PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

DEC 16 1929

DECEMBER 1929

Journal of Educational Research

Registered June 4, 1929

CONTENTS

		Page
The Purpose of a Standardised Test in Spelling	*Guy M. Wilson*	319
Knowledge of Success as a Motivating Influence in College Work	*E. C. Deputy*	327
The Diagnosis and Treatment of Some Common Difficulties in Solving Arithmetic Problems	*V. E. Chase*	335
The Effect of Specific Training on Ability to Read Historical Materials	*R. G. Simpson*	343
A Comparison of Column-Dictation and Sentence-Dictation Spelling with Respect to Acquisition of Meaning of Words	*H. W. Distad* and *Eva M. Davis*	352
Comparative School Marks of September and January Junior High-School Entrants	*O. M. Clem* and *J. K. Walklet*	360
On the Technique of Establishing the "Goodness" of an Intelligence Test	*Jack W. Dunlap*	367
Editorials		373
Reviews		376
Research Abstracts and Bibliographies		382
Research News and Communications		387

THE JOURNAL OF EDUCATIONAL RESEARCH is indexed in the Education Index

---

Copyright 1929, by Public School Publishing Co.

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879

JOURNAL OF EDUCATIONAL RESEARCH

EDITORIAL BOARD

E. J. ASHBAUGH
Dean, School of Education
Miami University,
Oxford, Ohio

A. S. BARR, *Chairman,*
University of Wisconsin,
Madison, Wisconsin

V. T. THAYER
Ethical Culture Schools,
33 Central Park West,
New York City

ASSOCIATE EDITORS

FRED C. AYER, *Professor of Educational Administration, University of Texas*

HARRY J. BAKER, *Clinical Psychologist, Detroit Public Schools*

FOWLER D. BROOKS, *Johns Hopkins University*

B. R. BUCKINGHAM, *Harvard University*

W. W. CHARTERS, *Director of Bureau of Educational Research, Ohio State University*

VIRGIL E. DICKSON, *University of California, Berkeley*

CHARLES FORDYCE, *Director, Bureau of Educational Measurements, University of Nebraska*

WILLIAM S. GRAY, *Dean, School of Education, University of Chicago*

CHARLES E. GREENE, *Director of Research, Denver Public Schools*

M. E. HAGGERTY, *Dean, College of Education, University of Minnesota*

V. A. C. HENMON, *Professor of Psychology, University of Wisconsin*

KARL J. HOLZINGER, *School of Education, University of Chicago*

ERNEST HORN, *College of Education, University of Iowa*

EARL HUDELSON, *Professor of Education, University of Minnesota*

ARTHUR W. KALLOM, *First Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston*

F. J. KELLY, *President, University of Idaho*

WILLIAM A. McCALL, *Associate Professor of Education, Teachers College, Columbia University*

EUGENE NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City*

W. J. OSBURN, *College of Education, Ohio State University*

SHELTON PHELPS, *Director of Instruction, George Peabody College for Teachers*

J. B. SEARS, *Stanford University, California*

J. L. STENQUIST, *Director, Bureau of Educational Research, Public Schools, Baltimore, Maryland*

GEORGE D. STRAYER, *Teachers College, Columbia University*

PERCIVAL M. SYMONDS, *Teachers College, Columbia University*

W. W. THEISEN, *Assistant Superintendent of Schools, Milwaukee, Wisconsin*

M. R. TRABUE, *Director of the Bureau of Educational Research, University of North Carolina*

W. L. UHL, *Dean, College of Education, University of Washington*

C. W. WASHBURNE, *Superintendent of Schools, Winnetka, Illinois*

GUY M. WHIPPLE, *Secretary, National Society for the Study of Education*

CLIFFORD WOODY, *Director, Bureau of Tests and Measurements, University of Michigan*

THE JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Company. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

Subscriptions and all communications about reprints or other business matters should be addressed to the PUBLIC SCHOOL PUBLISHING COMPANY, at Bloomington, Illinois.

Address all communications on editorial matters to A. S. Barr, Department of Education, University of Wisconsin, Madison, Wisconsin.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume XX DECEMBER 1929 Number 5

THE PURPOSE OF A STANDARDIZED TEST IN SPELLING

GUY M. WILSON
Boston University

MUCH of school work has become such a matter of habit that those performing it often fail to consider the fundamental purposes behind what they are doing. This has happened to all phases of school work including the giving of tests. It was the writer's privilege recently to go over the results of five years of testing in a nearby city. The superintendent considered the giving of tests a sign of progressiveness. These tests had been given, the papers had been scored, and the results had been filed. One test after another, as it had been advertised and brought to the superintendent's attention, had been used. Considerable time, labor, and expense had gone into this work, yet it was impossible to find that any constructive results had grown out of all this effort. Finally, the superintendent, himself, enrolled in classes in measurement, arranged for help by a graduate student who was working for a doctor's degree, and finally employed a competent assistant to help give tests which would transform this purposeless, *occasional testing* into a really *functional testing program.*

What are the purposes of testing? What are the qualities of a test that should be used to accomplish these purposes? In the first place, it should be recognized that a test is a tool. It is effective only, therefore, if it serves as a tool. The principal purposes are not in the giving of tests, but are in the results. If the school program is carried out for the aim of developing the child through instruction and worthwhile activities, then to have a part in such a school program, the test must be properly subordinated to this aim. The *first fundamental criterion* of a test, therefore, should be that it properly serves the main curricular aim of the subject being tested. The second fundamental criterion is that the test should properly reinforce good methods of teaching. There are many minor criteria relating to standardization, validity, etc., which have been, for the most part, well taken care of. It is the main criteria that are the ones which have been neglected.

It is proposed to analyze, herewith, a particular standardized test in spelling—The Stanford Achievement Test, Revised 1929. My attention was called to this test in its first form, since the evidence that the main purpose of teaching, the development of children, was apparently neglected in said test. When the revised form of this test appeared, I decided to give it immediate study. The results are shown below and speak for themselves. If a test is good it should be encouraged. If it is not good in its main bearings and influence, then that fact should be made known. A great deal of public funds have been spent for the testing movement, and it is time that we became critical of our standardized tests.

Table I shows for Form V of the Revised Stanford Achievement Test, the distribution of words for the eight grades, two to nine inclusive according to the Thorndike Word Book.[1] For instance, the 42 words in

TABLE I

SHOWING FOR FORM V, SPELLING, STANFORD ACHIEVEMENT TEST, REVISED 1929, THE DISTRIBUTION OF WORDS FOR EACH GRADE ACCORDING TO THE THORNDIKE WORD BOOK*

(Thorndike values are indicated at left)

	2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade	9th Grade	Total
1A	22	24	16	8	2	2	1	1	76
1B	8	10	10	8	6	2	2	2	48
2A	4	7	7	7	6	3			34
2B	3	4	9	9	8	6	5	5	49
3A	2	4	4	4	4	2	1	1	22
3B	1	2	3	6	6	5	5	5	33
4A		2	3	3	3	3	3	3	20
4B	1	1	2	2	3	3	3	3	18
5A		1	2	4	4	4	4	4	23
5B	1	1	1	2	2	3	3	3	16
6A		2	2	2	2	2	1	1	12
6B			1	1	1	2	2	2	9
7A			1	1	3	5	5	5	20
7B						2	2	2	6
8a-b		1	1	1	2	2	2	2	11
9a-b					1	2	2	2	7
10a-b					1	1	1	1	4
0		1	1	5	6	8	15	15	51
Total	42	60	63	63	60	57	57	57	459

* The four tables in this article were prepared under careful direction, by Dorothy M. Evans.

[1] Thorndike, E. L. *The Teachers Word Book.* (New York: Teachers College, Columbia University, 1921), 134 pp.

the test, which are to be used in Grade 2, are distributed according to the Thorndike Word Book from the first 500 (1A words) to the second part of the fifth thousand (5B words). Of the 60 third-grade words, 24 are in the first 500, 10 in the second 500, and thus the list continues on up through the second, third, fourth, and fifth thousand, and even beyond. One zero word occurs in the third-year list. Beginning with the fourth grade and thereafter, words are practically continuous through each grade from the first 500 words up to the zero words. The left-hand column in Table I indicates the value, e.g., "1A" means the first 500 of the first thousand; "1B" means the second 500 of the first thousand; "2A" means the first 500 of the second thousand; and so on.

The question at this point is how far up the list, using the Thorndike Word Book as a basis, should spelling words extend in order that the curriculum requirements will not be too easy or too difficult? It is not a simple matter to answer this question, but it is doubtful if the word list should in any case extend beyond the fifth thousand words. Such vocabulary studies as are available indicate that few grade children have common vocabularies in excess of 3000 words. There has been a disposition, therefore, to limit a grade spelling list to approximately 3000 words. The Boston Word List is considerably less. On this basis, it is doubtful if the second- and third-grade children should be given any words above the second thousand and many would confine the words to the first 1000.

The latter is doubtless more nearly correct. Beginning with the fourth grade, words might gradually be raised, but in no case, unless it be on the basis of individual choice, should words above the fifth thousand be given to grade children. This means that in every grade, beginning with the third grade, there are words in Form V of the Revised Stanford Achievement Test, which should not be regularly assigned to any grade child. The large number of zero words is quite inexcusable.

The common defense for extending words that are even beyond common adult usage is that in order to make a *valid* test words must be included which no child can spell. This is an axiom of the test maker which is doubtless applicable to research work only. The purpose of all work with children should be to teach them. If words are placed before them in their spelling tasks which are quite beyond them, the result is to confuse and discourage. This does not aid a proper teaching program.

It is possible at this point that there will be a definite disagreement between those who think mainly in terms of a curriculum adapted to children, and those who think mainly in terms of a testing program with

classification as the main end. These latter people have been chiefly responsible for promoting general survey tests and general classification tests. The value of such tests is very doubtful. Tests which conform to curricular standards and which serve good teaching are equally valuable for classification purposes when properly used. When reinforced by teachers' judgments and the right use of intelligence tests, the teaching test will serve any needed classification requirements.

An examination of the specific words of Form V will make its nature and construction clearer. It is very easy to approve for the second grade the simple 1A1 words such as *see, my, little, do, you, like, his,* and *your.* Most of the second-grade words are from the first one thousand. Twenty-six out of thirty-seven are so chosen. But this hardly justifies the inclusion of such a word as "agent" from the last part of the fourth thousand, or "banker" from the last part of the fifth thousand.

What is the purpose of spelling? Is it not to teach the child to spell correctly the words which he knows well enough to use regularly in written work? A generation or two ago the purpose was stated quite differently. It included building up vocabulary and disciplining the mind. We have learned, however, that the best discipline comes from useful tasks, that "cold storage" procedure is unprofitable, and that the right way to build up vocabulary is by extension of experience. Thus, our purpose on a functional basis becomes as stated above,—*To teach the child to spell correctly the words which he understands and wants to use often in his written work.* With this thought in mind, let us notice words of Form V used in other grades.

Grade 3 uses all of the second-grade words and adds twenty-two others. The words added include "region" from the second thousand, "area" from the third thousand, "reception" from the fifth thousand, "urgent" and "suggestion" from the sixth thousand, "stylish" from the eighth thousand, and "undefined" from the zero words. A zero word is above the tenth thousand and does not appear in the Thorndike Word Book. The question arises, 'Is a test containing such words for third grade a good test in spelling? If one answers that it is good he should be able to indicate for whom it is good. Certainly, it is not good for the children.

In the upper grades as indicated by Table I the difficulty increases rapidly, and we have in the fifth grade five zero words, in the sixth grade six, in the eighth grade fifteen, and in the ninth grade fifteen. There is in addition to this a generous distribution of words in all of these grades in

the fifth thousand and above. Fourth-grade children, for instance, are expected to understand the "theft of precious metal" and to "appreciate the friendliness of my successor." The fifth-grade child is expected to understand that a "committee of peasants was temporary," and to "eliminate disgusting controversies." Seventh-grade children are asked to spell, and, according to right curricular principles of spelling, are therefore expected to understand and use frequently in their written work such words as *plebiscite, dissension, prior, chancellor, rheumatism,* and *covenant.* Eighth-grade children are expected to frequently use such words as *phrenologist, penitentiary, poultice, rheostat, seismograph,* and *aeronautics.*

Looking over such lists of words for grade children one is reminded of the definition of a fanatic, given by Dr. Leonard Ayres a few years ago, as one who redoubles his efforts after forgetting his aims. Some of these words are beyond average adult age. In a summer session group of 70, one-half of whom were graduate students, only 20 knew the word *seismograph,* and only 16 the word *plebiscite.*

There are two forms of the revised Stanford Achievement Test, Form V and Form W. The spelling words for Form W are analyzed in

TABLE II

SHOWING FOR FORM W, SPELLING, STANFORD ACHIEVEMENT TEST, REVISED 1929, THE DISTRIBUTION OF WORDS FOR EACH GRADE ACCORDING TO THE THORNDIKE WORD BOOK

(Thorndike values are indicated at the left)

	2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade	9th Grade	Total
1A	23	24	15	7	4	1			74
1B	10	16	16	15	11	6	3	3	80
2A		5	7	8	8	9	6	6	49
2B	2	4	5	5	4	3	2	2	27
3A	2	2	4	6	5	5	5	5	34
3B	2	3	5	5	5	3	3	3	29
4A			1	2	2	2	2	2	11
4B		1	2	3	3	3	3	3	18
5A		2	2	2	2	2	1	1	12
5B	1	1	2	2	3	3	3	3	18
6A			1	2	4	4	4	4	19
6B					1	1	1	1	4
7A	1	1	1	1	1	1	1	1	8
7B			1	2	3	3	5	5	19
8a-b					1	1	1	1	4
9a-b				1	1	1	1	1	5
10a-b						3	2	2	7
0	1	1	1	2	2	6	14	14	41
Total	42	60	63	63	60	57	57	57	459

Table II. Analysis and explanation is similar to that for Table I. At the left is shown the Thorndike values from the first 500 (1A words) up to the zero words. The analysis is made for each grade separately, Grades 2 to 9. Most of the words are used in more than one grade, so that the total number of words in the test is not as great as the sum of the totals at the bottom. In other words, Table II shows for Form W, as does Table I for Form V, the analysis of words assigned to each grade regardless of the fact that some words may be assigned to more than one grade.

Table II shows that Form W is very similar in its general distribution of words to Form V. Zero words begin in Grade 2 and continue in every grade through the ninth. There is a general distribution of words throughout in the fifth thousand and above.

The Thorndike Word Book is a study of words based upon a combination of writing and literary vocabularies, but the literary basis is much stronger than the writing basis. Many, therefore, look upon the Thorndike Word Book as a reading vocabulary. The Commonwealth List of the ten thousand words most frequently used in correspondence[2] is based directly upon correspondence. Many regard it, therefore, as a better check for spelling lists.

Using the Commonwealth List as a basis and tabulating the results in a similar manner, we have Tables III and IV. Table III shows the spelling words for Form V of the Revised Stanford Achievement Test. The index numbers in Tables III and IV have been changed (as were those in Tables I and II) to indicate, by the numbers themselves, the thousands or part of the thousands in which each word falls. These index numbers are in the left-hand column. Careful comparison with Tables I and II will show greater agreement than differences. Comparing the totals at the right is possibly the easiest method of comparison. It will be noted that, checked by either list, there are a considerable number of words above the fifth thousand and enough of these in the ninth or tenth thousands or zero lists to justify considerable apprehension.

Table IV shows for Form W the distribution of the words on the Commonwealth List. This can be directly compared with the distribution on the Thorndike List by comparing Table IV with Table II. There are, of course, slight variations, but the general distribution is remarkably similar. Here again we find a considerable number of words above the

[2] Horn, Ernest. *A Basic Writing Vocabulary.* (Iowa City, Iowa: University of Iowa, 1926), 225 pp.

TABLE III

SHOWING FOR FORM V, SPELLING, STANFORD ACHIEVEMENT TEST, REVISED 1929, THE DISTRIBUTION OF WORDS FOR EACH GRADE ACCORDING TO THE COMMONWEALTH WORD LIST

(Commonwealth values are indicated at the left)

	2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade	9th Grade	Total
1A	19	22	16	9	5	5	4	4	84
1B	7	7	7	7	3	1	1	1	34
2A	3	5	4	4	3	2	1	1	23
2B	4	6	7	8	9	5	4	4	47
3A	3	5	6	5	4	4	3	3	33
3B	2	4	6	7	6	5	4	4	38
4A	1	3	3	4	4	3	3	3	24
4B	1	1	2	3	2	2	2	2	15
5A					3	4	4	4	15
5B	1	1	1	1	1	2	2	2	11
6		2	4	4	5	5	3	3	26
7		1	1	2	2	2	1	1	10
8				1	1	1	1	1	5
9	1	2	4	4	4	3	3	3	24
10			1	1	1	1	1	1	6
0		1	1	3	7	12	20	20	64
Total	42	60	63	63	60	57	57	57	459

TABLE IV

SHOWING FOR FORM W, SPELLING, STANFORD ACHIEVEMENT TEST, REVISED 1929, THE DISTRIBUTION OF WORDS FOR EACH GRADE ACCORDING TO THE COMMONWEALTH WORD LIST

(Commonwealth values are indicated at the left)

	2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade	9th Grade	Total
1A	18	20	11	6	4	2	1	1	63
1B	7	9	10	9	6	4	3	3	51
2A	3	7	9	9	9	8	6	6	57
2B	4	5	9	9	9	9	8	8	61
3A	2	5	6	6	6	4	2	2	33
3B	7	8	10	11	8	6	6	6	62
4A		2	2	2	2	2	1	1	12
4B	1	2	3	2	2	3	2	2	17
5A				1	1	2	2	2	8
5B		1	1	2	2	2	2	2	12
6a-b					2	3	3	3	11
7a-b				2	3	3	3	3	14
8a-b				1	2	2	2	2	9
9a-b					1	1	1	1	4
10a-b			1	1	1	1	1	1	6
0		1	1	2	2	5	14	14	39
Total	42	60	63	63	60	57	57	57	459

fifth thousand limit and the largest number of these at any single point in grades six to nine comes at the zero list or above the ten thousand limit.

It is possible that my study of spelling has been too limited to permit me to judge adequately of the spelling test which I have attempted to criticise. The work gives evidence of unusual statistical care. The minor criteria applicable to the formation of a test have been fully heeded. There is a growing tendency, however, to regard the minor criteria as quite secondary to the major criteria, and to regard the whole testing movement as a tool quite subordinate to methods of teaching and the curriculum. If testing is regarded, for instance, as the major consideration and the real master of the whole teaching process, then we will be likely to continue, as shown by this test, to bring into the schools meaningless tasks, formal procedures, and a reliance upon the mechanics of statistics and tests to do what cannot be done on the basis of humane and sensible treatment of children.

Without going into detail, I must insist that children can be classified without the use of such unreasonable material. The right test for spelling, if the word list is properly made up and adjusted to the ability of the child, is his entire word list. Spelling is a drill subject. The procedure is a drill procedure. It waits only upon the understanding of the child, the feeling of need for the particular words, and the right adjustment of curricular load to the ability of the child; and then when all of that is done, the child should be expected to master every word assigned, to check on any word missed, to carry that word forward and at the end of a stated period to spell correctly every word in his list. This, it seems to me, makes superfluous any extended testing-classification procedure such as is set up in this testing program for spelling. A testing program which rides over curricular principles and right methods of teaching as ruthlessly as do the spelling tests here analyzed, should lead administrators to pause for their bearings. These statements are made carefully, thoughtfully, with the child and the public funds immediately in mind. But in the long run, our profession itself is concerned. We should be interested in putting over the right programs in the schools, and there should be enough courage among us to study and criticize new material presented. This should be done from a strictly professional standpoint, and for the best interests of all,—child, taxpayer, and profession.

KNOWLEDGE OF SUCCESS AS A MOTIVATING INFLUENCE IN COLLEGE WORK

E. C. DEPUTY
Teachers College, Columbia University

INTRODUCTION

THE influence of a knowledge of one's success in an activity upon his improvement in the activity has been studied and reported from time to time in the literature of Educational Psychology. Thus, Arps[1] studied the effects of knowledge of work done upon the score in lifting an ergographic load with the middle finger of the right hand until the subject was no longer able to budge the load. He concludes that, "the average unit value of the known series of all sets of experiments shows a higher rate of work by 18 percent over the corresponding value of the unknown series." Book and Norvell[2] found that the experimental group improved much more than the control group in making the small letter "a," cancellation, substitution, and mental multiplication when the former were asked to count and record their scores "and in other carefully planned ways were interested as much as possible in the improvement they were making." The original stimulus group suddenly ceased to gain when the incentive conditions were removed, while the control group which had been making little progress began suddenly to improve when the incentives were applied. Chapman and Feder[3] used incentives similar to those used by Book and Norvell with a beneficial effect. Spencer[4] who repeated Judd's 1905 experiment, having subjects place a dot upon what they took to be a true extension of a visible line behind a screen, finds improvement for those without knowledge of their success in three out of four cases.

Burnett and Pear[5] suggest that the problem of motivation is among

[1] Arps, George F. "A Preliminary Report on Work with Knowledge Versus Work without Knowledge of Results." *Psychological Review*, XXIV, 449-455.

[2] Book, W. F., and Norvell, L. "The Will to Learn: An Experimental Study in Incentives to Learning." *Pedagogical Seminary*, XXIX, 305-362.

[3] Chapman, J. C., and Feder, R. B. "The Effect of External Incentives on Improvement." *Journal of Educational Psychology*, VIII, 469-474.

[4] Spencer, L. T. "The Effects of Practice without Knowledge of Results." *American Journal of Psychology*, XXXIV, 107-111.

[5] Burnett, I., and Pear, T. H. "Motives in Acquiring Skill." *British Journal of Psychology*, XVI, 77-85.

the most important in the psychology of skill.

Several educational psychologists have mentioned the influence of a knowledge of success on improvement in school work. Thus, Starch[6] considers the knowledge of success and failure one of the potent factors in improvement in school subjects. Colvin[7] states that, "in the elementary and secondary schools, examinations, not as tests, but as incentives for keeping in mind the work that has been pursued over an extended period, should be emphasized more than they are at present." But Norsworthy and Whitley[8] do not think that the influence of the knowledge of one's success upon his performance can necessarily be transferred from the laboratory situation to the classroom. They point out that what is to be learned is much more definite under laboratory conditions, and that the law of effect plays its part immediately in the laboratory situation. Hence, there is no difficulty in making the connection between the satisfaction and discomfort and the bonds concerned.

PROBLEM

The problem of this investigation is: What influence will a frequent knowledge of success have upon the scores (grades) of the three groups of students in freshman philosophy?

PROCEDURE

In the spring semester of the academic year 1927-28 the writers taught three parallel sections of freshman philosophy at one of the state universities. All of the students were second-semester freshmen, except approximately ten percent who had entered the university directly from the high schools for the second semester. The course was elective and there were thirty-five students in each section. Storck's *Man and Civilization* was used as a text.

At the first meeting of the second week the three sections were given the Otis' Self-Administering Test of Mental Ability: Higher Examination with the thirty minute time limit. At the next meeting of the classes they were told, that from the instructor's past experience, he

[6] Starch, Daniel. *Educational Psychology.* (Revised Edition, New York: The Macmillan Co., 1927), p. 177.

[7] Colvin, S. S. *The Learning Process.* (New York: The Macmillan Co., 1913), p. 176.

[8] Norsworthy, Naomi, and Whitley, Mary T. *The Psychology of Childhood.* (New York: The Macmillan Co., 1925), pp. 197-199.

had found that students in the philosophy course were more successful when they were given brief written exercises to indicate their progress. Due to the amount of work entailed, however, in grading the papers he could not do this in all three sections at the same time. Accordingly, two sections would have the written exercises during the first half-semester and the third during the last half semester.

Section 1 which met on Tuesday and Thursday at nine o'clock was given ten minutes written work covering the preceding lesson each time the class met. Section 3 which met at one P.M. on Wednesday and Friday was given about twenty minutes written work on Friday covering the two preceding lessons. It was so arranged that one-half of the written work of Section 3 was the same as Section 1 had had on Thursday, but the other half was the written work that Section 1 would have the following Tuesday. Hence, the opportunity for the exchange of information outside of class was equalized. Section 2 which met on Monday and Wednesday at eleven A.M. was the control group for the first half-semester. The first ten minutes of each meeting in this section was devoted to reviewing, orally, the preceding lesson by discussing the points contained in the written work of Sections 1 and 3. The written work of the latter two sections was intended to serve as their review.

The lessons were conducted on the question and discussion basis, following the book rather closely. In order that the content of the lessons should be uniform in the three sections, the instructor prepared a definite outline for each assignment in which he indicated all points to be discussed and all illustrations to be used. The same outline was followed for each of the three sections.

The written exercises were all of the informal objective type, including multiple choice, completion, matching statements, and a few questions which asked the students to reproduce exactly certain topics or statements of the text. Not more than five percent of the questions were of this latter type. No so-called true-false statements were used. A graduate student assistant and the instructor scored both daily tests and examination papers independently. The scoring was sufficiently objective so that there was practically no disagreement. In case a student was absent, he was required to take the written exercises before his excuse card was signed.

The number of mimeographed copies of the exercises were carefully checked each time they were given to the students so as to prevent the circulation of any copies outside the class. It was explained to the stu-

dents that, with the objective type of question, it was absolutely necessary for each student to keep his eyes on his own work and that in case he did not, reductions would be made from his score. The instructor used as much tact as possible in presenting all the details of the procedure for the sake of the proper attitude toward the work on the part of the student. Evidence will be offered later that in at least one section, and possibly in two, such an attitude was maintained.

After discussion with the two sections which were taking the written class exercises, it was decided to let the daily written work count as two-thirds of their mid-semester grade.

The scores on the written exercises were put on the board by means of a code number at each class meeting immediately following the one at which the written exercises were done. The students were urged to keep a record of the total number of points earned, a thing which all but two or three did. The total scores to date for each student in Sections 1 and 3, were put on the board at three different equally distributed times the first half semester.

At the time of the mid-semester examination the experimental sections had had six weeks of the written and recitation work as described above. All three sections had the same type of review preceding the examination. The mid-semester examination was held at the regular class periods on two different days. Section 1 and the control section had the first half of the examination first, and Section 3 had the last half first so that the opportunity for exchange of information outside of class was equalized. Considering the fact that the questions were of the objective type, and that each half contained approximately one hundred items, it is doubtful whether any pupil's score was changed appreciably by such exchange of information.

TABLE I

THE MEAN INTELLIGENCE AND MEAN MID-SEMESTER EXAMINATION SCORES OF THE THREE SECTIONS IN FRESHMAN PHILOSOPHY

	Section	No. in Section	Mean Score	S.D.	Difference between Means and P.E. of Difference	Frequency of Written Work
Otis Self-Administering Test: Advanced	1	30	47.80 ± .91	7.38	$M_1 - M_3 =$ 2.94 ± 1.60	
Advanced	3	33	44.86 ± 1.33	11.34	$M_3 - M_2 =$ 1.23 ± 1.65	
Advanced	2	33	46.09 ± .99	8.40	$M_1 - M_2 =$ 1.71 ± 1.35	
Mid-semester Examination	1	30	118.40 ± 1.59	12.92	$M_1 - M_3 =$ 9.58 ± 2.29	Each Meeting
Examination	3	33	108.82 ± 1.66	14.16	$M_3 - M_2 =$ 3.49 ± 2.30	Once each Week
Examination	2	33	105.33 ± 1.60	13.64	$M_1 - M_2 =$ 13.07 ± 2.26	None except Mid-semester

It will be seen from Table I that, although the mean intelligence scores are not the same, the difference between any two means is not sufficiently large to be statistically significant—that is, four times the P.E. of the difference. The difference (2.94) between the mean of Section 1 (47.80), and the mean of Section 3 (44.86) is not twice as large as the P.E. of the difference (1.60). The P.E. of the difference (1.35) is not much smaller than the difference (1.71) between the mean of Section 1 (47.80) and the mean of Section 2 (46.09). When we examine the differences between the means of the mid-semester examination, however, we find that the difference (9.58) between M_1 (118.40) and M_3 (108.82) is more than four times the P.E. of the difference (2.29). The difference between M_3 and M_2 (3.49) is only about one and one-half times the P.E. of the difference (2.30). But the difference between M_1 and M_2 (13.07) is considerably more than four times the P.E. of the difference (2.26).

The section, which had written work each time it met, did significantly better than the one which had written work once each week or the one which had only the recitation work. But the section, which had written work only once each week, was not enough better, to be significant, than the one without any written work. Unless factors other than the examination are influencing Section 3, it seems that the interval between written work is too great to be as effective as it was in the case of Section 1.

The reliability coefficient of the Otis' Test of Mental Ability as quoted by Freeman[9] is .92 ± 2.5 points in 74. The Pearson Product-moment r between the total daily scores and the mid-semester examination grades for 66 cases was .79 ± .03.

During the last half of the semester, Section 2 became the experi-

TABLE II

THE MEAN FINAL EXAMINATION SCORES, AND MEAN INTELLIGENCE SCORES

	Section	No. in Section	Mean Score	S.D.	Difference between Means and P.E. of Difference	Frequency of Written Work
Otis Self-Administering* Test: Advanced	1	30	47.80 ± .91	7.38	$M_1 - M_3 = 3.20 \pm 1.65$	
Advanced	3	31	44.60 ± 1.38	11.40	$M_2 - M_3 = 2.60 \pm 1.73$	
Advanced	2	30	47.20 ± 1.40	8.43	$M_1 - M_2 = 0.60 \pm 1.38$	
Final Semester Examination	1	30	69.87 ± 1.38	11.24	$M_1 - M_3 = 1.68 \pm 2.00$	None except Final examination
Examination	3	31	68.19 ± 1.45	12.00	$M_2 - M_3 = 3.87 \pm 2.02$	None except Final examination
Examination	2	30	72.06 ± 1.41	11.44	$M_2 - M_1 = 2.19 \pm 1.97$	At each Class Meeting

* Three students had dropped from Section 2, and two from Section 3.

[9] Freeman, F. N. *Mental Tests.* (Boston: Houghton Mifflin Co., 1926), p. 184.

mental section, having written work at each meeting of the class, and Sections 1 and 3 became the control sections. Except for the change in sections, the work was conducted in every way similarly to the work of the first half semester.

The final examinations came in the following order: Section 3, Section 2, and Section 1. The questions were of the objective type as described above and were the same for the three sections. The students took all the time they desired to finish the examination, which consisted of approximately one hundred items.

TABLE III

THE PERCENTS OF SECTIONS 1 AND 3 ABOVE THE MEDIAN OF SECTION 2 IN THE INTELLIGENCE TEST, MID-SEMESTER EXAMINATION, AND FINAL EXAMINATION

	Section	On Otis Intelligence Test	On Mid-Semester Examination	Amount of Written Work	On Final Examination	Amount of Written Work
Percent above Median of Section 2	1	56%	86%	Each Meeting	46%	None
Median of Section 2	3	50%	54%	Alternate Meetings	58%	None

Table II shows that the difference (2.19) between M_2, the mean of the experimental group, and M_1 in the final examination is less than two times the P.E. of the difference (1.97), and the difference (3.87) between M_2 and M_3 is less than twice the P.E. of the difference (2.02).

Table III shows the percent of Sections 1 and 3 which were above the median of Section 2 in the intelligence test, mid-semester examination, and final examination. In Section 1 we find 56 percent above the median of Section 2 in intelligence, 86 percent in the mid-semester examination, and only 46 percent above in the final examination. In Section 3 we find 50 percent above the median of Section 2 in intelligence—in other words they are equivalent in intelligence—54 percent above in the mid-semester examination, and 58 percent above in the final examination. This 58 percent of Section 3 above the median of Section 2 is not as significant as it appears, for had the median score of Section 2 been one point higher, only 50 percent of Section 3 would have exceeded the median of Section 2.

The Pearson Product-moment r between the total daily scores and the final examination scores for 30 cases was $r = .66 \pm .07$.

The evidence from Tables II and III indicates that the knowledge of their daily success did not stimulate Section 2 to greater achievement as compared with the other sections on the final examination, a thing which it seemed to do for Section 1 on the mid-semester examination.

To throw further light on the problem the data in Table IV were collected.

TABLE IV

THE NUMBER OF HOURS PER WEEK OF CLASS WORK, OF OUTSIDE ACTIVITIES, AND THE PERCENT OF EACH SECTION PREFERRING DAILY WRITTEN WORK

	Average Number Hours per Week of School Work	Number of Hours per Week at Outside Activities	Percent Preferring Daily Written Work
Section 1	15.3	5.7	86
Section 3	14.9	8.85	85
Section 2	15.1	7.40	46

The second column of Table IV shows the average number of class hours per week taken by each section. There is no significant difference between the sections in this respect. Column three shows the average number of hours per section devoted to such extra-curricular activities as athletics, Y.M.C.A. and Y.W.C.A. work, debating, working for expenses, etc. Near the close of the semester the instructor asked the students in each section to record on paper which of the two methods—with or without written class exercises—they preferred. The report was made anonymously so that fear of making an unfavorable impression upon the instructor would not enter.

There is a rather striking difference between the attitudes of Sections 1 and 3 on the one hand, and Section 2 on the other hand. At three different times soon after Section 1 had been changed from an experimental to a control section, the students as a class asked to continue the daily written work. Although Section 3 expressed itself as favorable toward the written class work, their score in the mid-semester examination was not so gratifying as that of Section 1. There are at least two factors that may influence this: first, as Table IV shows, they are doing more outside work than either of the other sections, and second, their written work coming only once each week, there may have been a tendency to neglect the preparation of their class work more than in Section 1. It is suggestive to know that the extent of the unfavorable attitude of Section 2 toward the written exercises is due to the fact that their written work came during the second half of the semester, after a half semester of only oral recitation work.

SUMMARY

Many experiments concerning the influence of knowledge of success, upon the learning of skills, have indicated that learning takes place more rapidly when subjects know their success. When this principle was

applied to the classroom situation, its success with the three sections of students studied in this experiment seemed to depend upon the attitude that the students took toward the procedure and how frequently the students' progress was measured and made known to them. Certainly, in this experiment, we did not find the improvement in the control section (Section 2), when it became the experimental section, that Book and Norvell[10] observed in their control group when the material learned consisted in cancellation, substitution, and making the small letter "a." This suggests the timeliness of Henmon's[11] remarks concerning the precaution to be taken in deriving our principles of learning from studies in functions unlike those of the classroom situation, such as learning telegraphy, typewriting, ball-tossing, mirror-writing, and substitution experiments.

CONCLUSIONS

1. In the three sections of freshman philosophy studied in this experiment, the mean score was significantly increased when written exercises to measure the students' success were given each time the class met, provided the attitude on the part of the students toward the work was favorable.

2. The attitude of these students toward their work seemed to have a significant influence on their improvement even when their success was checked by written work each time the class met.

3. Considerable precaution should be taken in applying principles, derived from laboratory and other non-classroom situations, to work in school subjects.

---

[10] Book, W. F., and Norvell, L. "The Will to Learn: An Experimental Study in Incentives to Learning." *Pedagogical Seminary*, XXIX, 305-362.

[11] Henmon, V. A. C. "Needed Research in the Field of Learning." *Journal of Educational Research*, XI, (May, 1925), 313-321.

THE DIAGNOSIS AND TREATMENT OF SOME COMMON DIFFICULTIES IN SOLVING ARITHMETIC PROBLEMS

V. E. CHASE
Fordson Public Schools

SUCCESS in school has very commonly been predicated upon success in arithmetic. Failure in arithmetic has generally met with more concern upon the part of teachers and with more serious consequences to the pupil than has failure in reading, spelling, geography, or any of the other subjects of the curriculum.[1] This is probably due, in part, to the fact that when pupils leave school their knowledge of arithmetic and their ability to use it are more apt to become matters of immediate concern to their employers and others than is their knowledge of other subjects. The practical and social importance, which thus becomes attached to arithmetic, has made it the subject of many investigations. Most of these investigations have centered around an attempt to discover the kinds and nature of the mistakes responsible for failure and difficulty in the use of the fundamental processes and number facts.

The present investigation is concerned not so much with the number of mistakes which pupils make as it is with the reasons why they make them. What sort of synapse, mental tendency, or maladjustment is it which habitually produces error instead of accuracy?

Why, after years of training, presumably intended to establish correct habits in the use of fundamental processes and number facts, do many pupils persist in error and improper habits of reasoning? Such is the problem which confronts us in the present study. It is based upon the hypothesis that the mistakes and errors, which pupils are commonly observed to make, are merely symptoms and seldom, of themselves, constitute the true causes or mental maladjustments responsible for difficulty and failure. This hypothesis obviously is far-reaching. It implies that reasoning difficulties, like rheumatic pains, may sometimes be due to causes quite remote from the actual point where the trouble is manifested. It also implies that the meager and disappointing results, sometimes experienced by teachers after prolonged drill for corrective purposes,

[1] "'. . . this subject (arithmetic) is the chief source of non-promotion in the elementary school . . .'"—Buswell, Guy T., and Judd, Chas. H. *Summary of Educational Investigations Relating to Arithmetic.* (Chicago: University of Chicago, 1925), p. 7.

may be due to the fact that the drill methods and material employed have been in the nature of local treatment, as it were, and have not actually reached the seat or origin of the difficulty at all.

The practical significance of this hypothesis was pointed out by Woody in 1916 when he called attention to the need for a type of diagnosis which would aim to reveal the mental processes which cause error.[2] This is a very important point, because it differentiates between the process of finding and listing types of errors and the process of ascertaining the mental maladjustments which are responsible for such errors; it shows the need not only for a knowledge of symptoms but for a means of ascertaining their causes; it points out the way to a technique of diagnosis which should eventually bring systematic, scientific, and corrective treatment much nearer to the classroom.

The following investigation was undertaken with a full appreciation of the magnitude of the problem, but it was also undertaken with the conviction that future progress in the direction of corrective and preventive instruction must take into account the value and possibilities of the case-study method and individual diagnosis as a means for discovering and adequately treating *causes* of learning difficulties. It was felt that any attempt to thus investigate further the causes of difficulty in solving reasoning problems, by a diagnosis of individual cases of maladjustment, and to contribute something to the present technique of remedial procedure should not be without some profit to both teachers and pupils.

SELECTION AND PRELIMINARY EXAMINATION OF CASES

Teachers of the Fordson Junior High Schools were asked in the fall of 1927 to report upon blanks, provided for the purpose, all pupils who were doing unsatisfactory work in arithmetic and who were approximately normal in intelligence or even above. The teachers were also asked to state in this initial report what, in their judgment, appeared to be the most outstanding weakness or cause of difficulty and to give the Stanford reading and arithmetic scores for October 1927. A total of fifty-four initial reports were received, from which seventeen cases were selected for intensive study and observation. The IQ of this group ranged from 84 to 118.

---

[2] Woody, Clifford. "Measurements of Some Achievements in Arithmetic," *School and Society*, IV, (Aug. 19, 1916), 299-303.

For each of the seventeen individuals selected, a detailed case history was prepared by the classroom or home-room teacher upon forms provided for the purpose. This case history included:

1. Observed symptoms of disability
2. Nativity
3. Home Conditions
4. School History
5. Arithmetic History
6. Character and Temperament
7. Physical History

The information contained in these individual case histories later proved to be of frequent aid for diagnostic purposes and for the determination of remedial procedure. After an examination of the detailed history of each case had been made, a battery of tests in the fundamentals and in reasoning was given to ascertain as many specific causes of difficulty as possible.

DETERMINING THE CAUSES OF REASONING DIFFICULTY

An analysis of the errors produced in the above mentioned tests, was made. The element or elements in the situation which seemed to cause more or less significant or persistent errors were isolated. These elements were then incorporated into verbal problems which were applied in a further intensive study of the subjects' reasoning difficulties in the following manner:

The subject was placed in as congenial an atmosphere as possible, with the examiner and a stenographer present, and asked to solve the problems aloud using pencil and paper when necessary. A complete stenographic record was kept of all conversation and statements made by both the pupil and the examiner. In this examination the examiner attempted at all times to avoid leading questions. He tried only to keep the pupil's attention fixed upon the problem in hand in order that the habitual mental processes of the pupil might have free play and not be inhibited or disturbed in any manner.

This examination provided the investigator with, (a) a complete stenographic record of the pupil's reactions to verbal and typically difficult problems, (b) an objective record of the examiner's close-range observations of the character of the pupil's activity in solving problems, and, (c) the pupil's own written record of solutions. This constituted the material from which an analysis of reasoning difficulties was made, a diagnosis projected, and the remedial treatment ascertained. A sum-

mary statement of the procedure and the results as outlined above was then made for each case, and a proposed course of remedial treatment was applied in special classes to which the cases were assigned.

THE CASE OF CORNELL P.

A typical case is that of Cornell P., who was reported as failing in arithmetic, because (according to his teacher) of a lack of interest and unsatisfactory work. His chronological age was twelve years and six months. He possessed an IQ of 98, showed a grade level of 7-4 in reading, and 4-7 in arithmetic. He was in grade 8A at the time he was reported. The case history showed that Cornell was born of Roumanian parents in the United States. The family had changed homes ten times. During the six years that Cornell had been attending school he had skipped kindergarten, and grades 1B, 3A, and 5B.

In character and temperament,[3] he rated highest in the control qualities of honesty and truthfulness; low in the guidance qualities of initiative, inventiveness and discernment; low in the motive qualities of interest and ambition; and high in vitality. At the time he was reported, he was doing a D grade of work in arithmetic and a C grade of work in literature, reading, and shop. He appeared to possess a particular aversion to arithmetic.

Reaction to Tests.—The Wisconsin Inventory Tests in Arithmetic, the Stevenson Problem Analysis Test II, and the Sangren-Woody Reading Test were administered. His performance upon these tests ranged in grade levels from 4-0 to 6-1 in the various arithmetic abilities and from 5-3 to 9-0 in the various reading abilities measured.

An analysis of his errors indicated that among the important contributing factors, responsible for arithmetic difficulty in his case, were carelessness and a confusion of processes. For example he persistently made such errors as: 3 plus 3 equals 9, and 4×8 equals 12.

Diagnosis of Reasoning Difficulties.—While a lack of mastery of the combinations and fundamental processes was thus revealed as a factor to be considered in an attempt to correct the arithmetic disabilities of this pupil, it was felt that a further search would reveal other causes of difficulty more remote, perhaps, but equally if not more significant.

Therefore, instead of immediately setting up drill exercises to correct the errors with combinations and processes these errors were made the

[3] Terminology from Character Education Institution Circular under date of Jan. 1, 1926. Washington, D. C.

basis for further study and investigation. They were first incorporated into a series of simple problems similar to the following:

> "A family uses two quarts of milk per day. What does it cost this family per week for milk when milk is 12 cents per quart?"

These problems were then presented to the pupil in the manner described and a careful examination of his oral and written responses was made. This examination revealed the following additional facts regarding the causes of this pupil's difficulty in solving verbal problems:

1. Carelessness;
2. Inability to differentiate between linear, square, and cubic measure;
3. Little conception of quantitative relationships;
4. No conception of simple fractional relationships;
5. A lack of comprehension probably attributable to insufficient training in silent reading;
6. Failure to understand many of the arithmetical terms and concepts with which he was expected to be familiar; and
7. Failure to make up essentials in the arithmetic course of study which had been skipped as a result of double promotions.

It is obvious, of course, that some of the causes of difficulty in this case are attributable, not so much to the pupil's inability or failure to adjust himself to normal instructional situations as it is to the school's failure to adequately compensate for irregularities and conditions of its own making. For instance, if pupils are to be permitted to skip grades some provision should be made whereby they will not be deprived of the absolute essentials of the grade or grades skipped.

Remedial Treatment.—A detailed course of remedial treatment, aimed to correct the deficiencies in fundamentals and reasoning noted above, was set up and followed for a period of ninety days. The accompanying graph shows that during this period the pupil made an arithmetic growth of approximately one year as measured by the Stanford Arithmetic Test and a growth in specific arithmetic abilities varying from zero to over three years.

Remedial treatment was planned for each individual, the period of remedial instruction covering approximately three months. While the process was interrupted by various unavoidable circumstances, the results of the final tests show much improvement, some cases having made more than three years gain during the period in certain specific abilities in

which they were deficient. A gain from one to two years was common. Instances were reported by teachers of pupils who returned to their classes with renewed interest in arithmetic and who carried on a very satisfactory grade of work thereafter. One pupil reported that she "liked to come to remedial class because she could see that she was getting something from it."

While the causes of difficulty in solving verbal problems appeared to vary considerably in the different cases studied, all of them showed similar marked improvement in response to the individual remedial treatment set up.

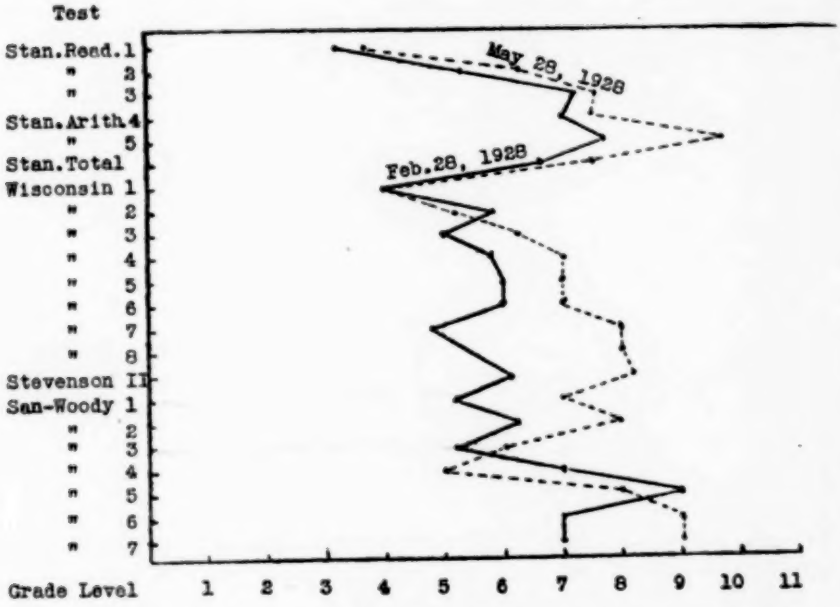


DIAGRAM NO. 1. SHOWING HOW A PUPIL WHO WAS FAILING IMPROVED IN MANY ABILITIES AFTER THREE MONTHS OF CORRECTIVE TREATMENT

CONCLUSIONS

1. A common cause of difficulty in solving verbal problems was found to be insufficient mastery of the fundamentals. Failure with the combinations involving 7, 8, and 9 were frequent. Nine of the twelve cases invariably failed on division problems involving zero quotients. Of the fundamental processes, division presented the most difficulty, being a serious cause of error in ten of the twelve cases.

2. Half of the cases could not tell which process to apply in solving verbal problems. Some focused attention upon the integral expressions in the problem; others applied the trial and error method until an answer was obtained which to them was satisfactory, but which often was incongruous; and, still others had the habit of doing problems in "sets," i.e., assuming all contiguous problems to be alike and solving them all according to the same "formula."

3. It is evident from this experiment, however, that lack of familiarity with combinations and processes is but one factor contributing to failure and not necessarily the most important. Other causes, such as improper reasoning habits, faulty reading and weak comprehension, were found to be responsible for much difficulty and failure. In three cases difficulty was due to rapid, rote reading; others experienced serious difficulty by reason of an inadequate vocabulary and a lack of knowledge of the meaning of words common to arithmetic. Following is an illustration from Case No. 8.

E. "What are you doing?"
S. "Subtracting."
E. "Why don't you add?"
S. "It doesn't say add, it says increase."
E. "What does 'increase' mean?"
S. "It means less."

In nine cases, failure was due largely to inability to sense relationships usually quantitative but sometimes grammatical. Five cases lacked the fundamental concepts necessary for comprehending and solving ordinary verbal problems. This appears to be due in some cases to the skipping of grades or double promotions, one of these cases having been double-promoted three times. Five lacked ability to retain the elements of a problem in mind while attending to new elements. The following quotation illustrates the nature of this difficulty:

E. "Have you found out how much he makes per day?"
S. "505.26."
E. "How much does he make per week?"
S. "That (pointing to $505.26) is how much he makes per week."

Four were in the habit of making random, premature attacks upon problems. The following from Case No. 3 typifies this difficulty:

E. "What does the problem call for?"
S. "How much it cost per loaf."
E. "Does it? Read the problem again."
S. (Reads problem again.)
S. "How much he made per day."

Additional causes of failure observed were: lack of development and drill in the correlated processes of analysis and synthesis; lack of training in estimating answers; lack of organizing power or ability to do consecutive thinking; lack of capacity or training in inductive reasoning; short attention span; and, lack of initiative and self-confidence.

4. It is apparent from this experiment that marked improvement can be brought about by the application of corrective treatment based upon a systematic study and diagnosis of individual difficulties. It is probable that many of the difficulties, discovered in this manner, may be found to be fairly common within a system and, therefore, much of the corrective treatment and technique evolved might profitably be incorporated into the instructional methods and materials of the schools.

THE EFFECT OF SPECIFIC TRAINING ON ABILITY TO READ HISTORICAL MATERIALS

ROBERT GILKEY SIMPSON*
Carnegie Institute of Technology

THIS study represents an attempt to find out experimentally what effect training in answering questions, evaluating, outlining, and summarizing has upon pupils' ability to read and study historical materials. The experimental technique suggested by W. A. McCall[1] in his book, *How to Experiment in Education*, was employed in the investigation.

A review of the previous contributions to directed study discloses the fact that much still remains to be done in this field. Various methods, both general and specific in nature, have been tested in an experimental way, from time to time, with considerable success. But as far as the writer knows no one has tested the effect of training in a group of specific skills which are believed to be essential in the process of learning, or study.

THE GENERAL PLAN OF EXPERIMENT

1. *Scope.*—There were 1,074 fifth, sixth, and seventh-grade pupils who participated in the experiment. Of this number, 574 pupils represented seven different school systems in Allegheny County exclusive of Pittsburgh; the remainder were from platoon schools in the city of Pittsburgh. The scores of only 606 pupils were finally included in the study since a large number of scores were eliminated in the process of pairing[2] groups.

2. *Selecting and Equating Groups.*—The method used was to select, in each school participating, at least two groups, the one being designated as the experimental and the other as the control group. In a number of instances several groups of a single grade were obtained. This afforded an opportunity to check on the different orders of rotation of the four new factors used in the special training.

---

* A portion of a Doctor's Thesis directed by G. A. Yoakam, University of Pittsburgh.

[1] McCall, W. A. *How to Experiment in Education.* (New York: Macmillan Co., 1926), Chapters VII and VIII.

[2] *Ibid.*, Chapters VII and VIII.

The following four criteria were employed in determining equivalent or parallel groups: the scholastic standing of each pupil in reading; the scholastic standing of each pupil in history for the previous semester; his intelligence quotient; and his standing in one of the initial tests. The records and scores made in the initial tests and school records were reduced to a comparable basis by the Russell technique[3] for finding M-scale scores.

3. *Control of Certain Variable Factors.*—In so far as possible the same instructor had charge of one or more experimental groups, and also a parallel[4] control group. The "practice effect" of training was checked on by the scheme of rotation of the experimental factors. This was accomplished by running a check group while the same teacher taught five groups, one of which was a control on the other four.

The equivalent group plan was checked against an identical[5] group arrangement to find out the accuracy of the former as a technique upon which to calculate the results. A comparison of the two plans gave evidence of the fact that the equivalent group procedure was quite satisfactory, since the mean gains and losses of each type of group compared corresponded almost identically with those of its control group.

In order to check the reliability of certain devised history tests, the writer correlated them with the Thorndike-McCall Reading and the Van Wagenen History Tests, which were run parallel with the non-standardized tests throughout the experiment. Since equivalent forms of certain tests could not be obtained in every case, the same form was employed. The repetition of the same tests probably had some influence upon the results of the experiment. But, according to Monroe,[6] the increased gain resulting from the repetition of the test probably does not exceed ten percent.

4. *Time Control.*—The regular class periods were thirty-five minutes in length, so divided that ten minutes were devoted to directed study, and ten minutes to class discussion. The remaining fifteen minutes were then given to the conventional method of conducting the recitation without any special effort to direct pupils' study. It was suggested that this

[3] Russell, Charles. *Classroom Tests.* Boston: Ginn and Company, 1926.

[4] Parallel, and equivalent, are used synonymously throughout the study.

[5] Identical groups were those which had the same, or identical M-scale score, pupil for pupil.

[6] Monroe, W. S. *An Introduction to the Theory of Educational Measurements.* (Boston: Houghton Mifflin Company), 167-168.

part of the recitation be conducted as the regular period for the control groups.

5. *Plan of Instructing Groups.*—The county experimental groups were given special training, according to instruction sheets placed in the hands of teachers, in answering questions, evaluating, outlining, and summarizing historical materials for periods of five recitations each; and the city experimental groups were given special training in these four factors in different orders of rotation for periods of eight[7] recitations each. This gave a total training period of one month for the county groups, and a period varying from two months, when classes recited daily, to about three months, when classes alternated with other school subjects for the city groups.

6. *Tests Used.*—Both the experimental and control groups were tested before and after practice. In the county, the groups were tested before and after a period of one month, and in the city the groups were tested before and after each period of eight recitations and the total period of thirty-two recitations. Forms A and B of the Van Wagenen History Tests, a specific fact history test, and a composite history test were used to check on the county groups. The Thorndike-McCall Reading Test, Forms 3 and 5, a special fact history test, and a historical reading ability test of four parts were used to check on the city groups. The two former tests were administered at the beginning and end of the total period of training, and different parts of the historical reading ability test were given at the beginning and end of each period of eight recitations. This shifting of the testing program made it possible to check specifically on the experimental factors used in the study.

THE RESULTS OF THE EXPERIMENT IN CERTAIN SCHOOLS OF ALLEGHENY COUNTY

The results given in the following three tables are based on records of 278 seventh-grade pupils representing seven different school systems in Allegheny County. The data are explained on the basis of the experimental coefficients[8] which McCall devised to interpret the significance of the difference between the experimental and control procedures. The writer used Yule's formula[9] for combining the distributions. The order

[7] Eight recitations were considered nearly equivalent to three weeks.

[8] McCall, W. A. *op. cit.*, Chapter VIII.

[9] Yule, G. U. *An Introduction of the Theory of Statistics.* (London: Charles Griffen Company), Chapter VIII.

of rotation of the experimental factors used in the county schools was answering questions, evaluating, outlining, and summarizing.

TABLE I

SUMMARY OF THE RESULTS OF THE EXPERIMENT IN THE ALLEGHENY COUNTY SCHOOLS AS MEASURED BY THE GENERAL UNITED STATES HISTORY TEST*

Experimental Method		Control Method	
$Mean_1$	31.20	$Mean_2$	31.68
$Mean_3$	39.36	$Mean_4$	40.27
SD_1	6.70	SD_2	5.58
SD_3	8.14	SD_4	9.22
SDm_1	0.67	SDm_2	0.56
SDm_3	0.81	SDm_4	0.92

Difference 1.39 SD_{31} 1.05 SD_{32} 1.08
SDd 1.50 P.E. of D. 1.01 E.C. —0.39

* This test was not an intrinsic part of the study.

The experimental coefficient for Table I is negative, and indicates that the chances are about six to one that there is a true superiority of the control procedure. It supports the inference that the special training was not effective in fixing historical facts in the minds of the pupils as measured by this test.

TABLE II

SUMMARY OF THE RESULTS OF THE EXPERIMENT IN THE ALLEGHENY COUNTY SCHOOLS AS MEASURED BY THE COMPOSITE HISTORY READING TEST

Experimental Method		Control Method	
$Mean_1$	67.52	$Mean_2$	74.95
$Mean_3$	93.80	$Mean_4$	85.29
SD_1	28.96	SD_2	32.70
SD_3	36.62	SD_4	31.93
SDm_1	2.45	SDm_2	2.77
SDm_3	3.16	SDm_4	2.70

Difference 15.94 SD_{31} 1.27 SD_{32} 1.22
SDd 1.76 P.E. of D. 1.20 E.C. 3.29

The experimental coefficient resulting from the composite history reading test as shown in Table II is 3.29. This is an exceptionally large coefficient and points to the fact that the chances are much greater than 65,000 to 1 that there is a true superiority of the experimental procedure. Stated in terms of practical certainty[10] we can be more than doubly certain that there is a true difference between the two groups in favor of the experimental technique. The inference to be deduced here is that the special training was effective in improving the abilities required in the composite history reading test.

[10] Practical certainty, according to McCall, signifies that the true M_1 lies between the obtained M_1 minus $3SDm_1$ and the obtained M_1 plus $3SDm_1$.

TABLE III

SUMMARY OF THE RESULTS OF THE EXPERIMENT IN THE ALLEGHENY COUNTY SCHOOLS AS MEASURED BY THE VAN WAGENEN HISTORY TEST

Experimental Method		Control Method	
$Mean_1$	70.28	$Mean_2$	72.42
$Mean_3$	72.47	$Mean_4$	72.43
SD_1	9.87	SD_2	9.76
SD_3	8.09	SD_4	10.94
SDm_1	0.94	SDm_2	0.93
SDm_3	0.76	SDm_4	1.04

Difference 2.58 SD_{s1} 1.20 SD_{s2} 1.39
SDd 1.83 P.E. of D. 1.23 E.C. 0.51

Table III gives an experimental coefficient of 0.51 for the Van Wagenen History Test. It is positive and fairly significant. Since an experimental coefficient of 1.0 means practical certainty, a coefficient of 0.51 would suggest that one can be only half certain that there is a true difference between the two procedures. In the case of this coefficient the advantage is in favor of the experimental method. The chances are greater than 11 to 1 that the special training was more effective than the usual classroom procedure in increasing those abilities required in the Van Wagenen History Test.

THE RESULTS OF THE EXPERIMENT IN THREE PLATOON SCHOOLS IN THE CITY OF PITTSBURGH

In many respects the experiment in the city schools duplicated that in the county. The same study material was used and the same general procedure was followed. There were, however, important differences. The special training factors were rotated in different orders and were extended over a longer period of time. The real purpose of the rotation scheme, as has been mentioned, was to check the effect of practice. Another significant purpose of the plan, was to note the effectiveness of the different orders of rotation on the pupils' ability to study history. The results of this phase of the study were compiled for 328 intermediate-grade pupils, of whom only 36 were in the fifth grade.

The order of rotation followed by five equivalent groups of sixth-grade pupils of the Conroy Platoon School was: for Group A, answering questions, evaluating, outlining, and summarizing; for Group B, evaluating, outlining, summarizing, and answering questions; for Group C, outlining, summarizing, answering questions, and evaluating; and for Group D, summarizing, answering questions, evaluating, and outlining. The order of rotation followed by the various other city groups of pupils

was similar to that indicated above. The results obtained for all these groups made it possible to check fairly accurately on the rotation plan employed throughout the experiment. Table IV gives the results for the equivalent groups.

TABLE IV

A SUMMARY OF THE EXPERIMENTAL COEFFICIENTS FOR FOUR EQUIVALENT GROUPS OF SIXTH-GRADE PUPILS VERSUS THE SAME PARALLEL CONTROL GROUP

New Factors	Experimental Method				Total Av. E.C.'s
	Group A E.C.'s	Group B E.C.'s	Group C E.C.'s	Group D E.C.'s	
Answering questions	PERIOD 1 0.73	Period 4 1.31	Period 3 0.64	Period 2 0.77	0.86
Evaluating	Period 2 0.82	PERIOD 1 1.86	Period 4 0.46	Period 3 —0.04	0.77
Outlining	Period 3 0.94	Period 2 0.40	PERIOD 1 2.62	Period 4 1.50	1.37
Summarizing	Period 4 0.57	Period 3 0.58	Period 2 0.14	PERIOD 1 0.88	0.54
Total Av. E.C.'s	0.77	1.04	0.97	0.77	

The words in large type in this table show the experimental coefficients resulting from the first periods of training. A comparison of these coefficients with those for other periods of training will show to what extent the effect of practice had on the experiment. The average experimental coefficients, in the column at the extreme right of the above table, indicate that outlining as a method of teaching was the more effective and that summarizing was the least effective method of the four. The average coefficients at the foot of the table indicate that rotating the experimental factors in the order of evaluating, outlining, summarizing and answering questions, gives a greater effect than rotating them in any other order. Rotating these experimental factors in the seemingly logical order—answering questions, evaluating, outlining, and summarizing—produces the least effect of all.

An inference that may be deduced from this table is that the effect of practice was not a predominating element in influencing the results of the experiment. Another inference, that is supported by the data in this table, is that when pupils are first taught to organize their materials in the form of an outline and then to answer questions about the context, their reading ability seems to be improved much more than when they are taught these special methods in the reverse order. In addition, it

seems obvious that the rotation order, evaluating, outlining, summarizing, and answering questions, is a more effective plan of teaching than any other order of rotation followed in this study.

TABLE V

SUMMARY OF THE RESULTS OF THE EXPERIMENT WITH FIVE SIXTH-GRADE GROUPS OF THE CONROY PLATOON SCHOOL AS MEASURED BY THE THORNDIKE-MCCALL READING TEST

Experimental Method		Control Method	
$Mean_1$	52.60	$Mean_2$	55.95
$Mean_3$	60.73	$Mean_4$	59.57
SD_1	7.12	SD_2	7.32
SD_3	9.06	SD_4	6.87
SDm_1	1.59	SDm_2	1.64
SDm_3	2.03	SDm_4	1.54

Difference 3.93 SD_{s1} 2.59 SD_{s2} 2.25
SDd 3.47 P.E. of D. 2.33 E.C. 0.40

The experimental coefficient for Table V is not large, but it clearly indicates a superiority of the experimental over the control procedure as measured by the Thorndike-McCall Reading Test. The chances are about seven to one that the difference is significant and warrant the inference that the difference is probably real.

TABLE VI

SUMMARY OF THE RESULTS OF THE EXPERIMENT WITH THREE SEVENTH-GRADE GROUPS OF THE CONROY PLATOON SCHOOL AS MEASURED BY THE THORNDIKE-MCCALL READING TEST

Experimental Method		Control Method	
$Mean_1$	50.71	$Mean_2$	52.43
$Mean_3$	61.50	$Mean_4$	55.09
SD_1	6.60	SD_2	6.82
SD_3	10.92	SD_4	8.00
SDm_1	1.46	SDm_2	1.45
SDm_3	2.33	SDm_4	1.70

Difference 8.13 SD_{s1} 2.72 SD_{s2} 2.23
SDd 3.52 P.E. of D. 2.37 E.C. 0.84

Table VI gives an experimental coefficient of 0.84, which supports the inference that there was a true difference between the two methods in favor of the experimental procedure. Expressed in terms of chances, this coefficient indicates that the chances are about seventy-five to one that the experimental method was more effective than the control method as measured by the Thorndike-McCall Reading Test.

The tables in which the results are recorded for the Beltzhoover and the James Rogers Platoon Schools are omitted in this abstract of the experiment since, in practically every instance, the inferences deduced from the foregoing tables are supported.

CONCLUSIONS

Since there are always many variable factors to be considered in an experimental study of this type, one needs to be reluctant in attempting to make sweeping conclusions. It is highly probable, however, that the data collected in this study would warrant the statement of certain definite trends in teaching pupils to read and study more advantageously.

1. If we define as a phase of organization the ability to answer questions, evaluate, outline, and summarize, then it must follow that pupils' ability to organize was improved by the experimental procedure. Again, if we define study, in part, as the ability to organize in the manner stated above, then it must follow also that pupils' ability to study was improved by the special training given in the experiment.

2. The study shows that certain orders of rotation of the experimental factors were more significant than others. For instance, the order of outlining, summarizing, answering questions, and evaluating, was a more effective means of reading, as measured by the Thorndike-McCall Reading Test, than any other order of rotation. The order of rotation beginning with evaluating was also shown to be an effective means of study. These facts seem to warrant the conclusion that if intermediate grades are taught to organize their materials of study before attempting to answer questions about the context, their ability to read will be improved considerably more than if they pursue their reading and study in the reverse order.

3. There is evidence of the fact that outlining, as a method of teaching, is more effective than any other special method employed in this experiment.

4. The results of the Van Wagenen History Test suggest that the rotation order of answering questions, evaluating, outlining, and summarizing was a more valuable method of teaching than the usual classroom procedure.

5. It seems obvious that the special training was not valuable in improving pupils' ability to retain facts as measured by the history test used for that purpose. This is, however, what might have been expected since certain definite units of subject matter could not be strictly adhered to.

6. Personal observation of the study convinces the writer, and those who were closely associated with the experiment from time to time, that the satisfaction resulting from the study exercises employed gives greater zeal and zest to school work.

IMPLICATIONS

The following implications seem more or less obvious:

1. The fact that there were a number of negative and small experimental coefficients might imply that periods of training not longer than two, four, or six weeks are far too short to produce any change in ability to read historical materials which have been previously learned in school.

2. Definite units of subject matter should be selected and submitted to training exercises similar to those used in this experiment. Then a fact and interpretation test based on these units should be devised and given to check the value of such a new procedure in fixing facts in the minds of pupils.

3. Material which children are to read should be presented so that they will be compelled to read the context before attempting to answer questions and to organize it according to a model plan.

A COMPARISON OF COLUMN-DICTATION AND SENTENCE-DICTATION SPELLING WITH RESPECT TO ACQUISITION OF MEANING OF WORDS

H. W. DISTAD
New Mexico State Teachers College
and
EVA M. DAVIS
Houston Public Schools

A NUMBER of studies comparing the efficiency of column or list spelling with context spelling have been reported.[1] All of these studies show a loss in efficiency through transfer. That is, when the same words are presented in both column and context forms, the percentage of correct spelling is greater with the former than with the latter method. Further, it appears that pupils who study or are tested on words in column form are at no disadvantage when the words appear in new context form. For example, Winch found that pupils who studied words in column form possessed greater ability to transfer these words into cursive dictation than did pupils who studied the same words in context form.[2] Hawley and Gallup conclude from the results of their experiment "that in general the loss in spelling ability which results when pupils write in context cannot be prevented any more effectively by using a sentence method than it can by using the list method."[3] A recent study by McKee shows that pupils who studied and were tested by the column form method were as efficient or more efficient than pupils who studied and were tested by the phrase, sentence, or paragraph forms when the words were presented in new context forms.[4]

---

[1] Wallin, J. E. *Spelling Efficiency.* (Baltimore: Warwick and York, 1911), pp. 65-76.
Cook, W. E., and O'Shea, M. V. *The Child and His Spelling.* (Indianapolis: Bobbs-Merrill, 1914), Ch. V.
Ashbaugh, E. J. *Comparison of Spelling in List and Context Form.* (Unpublished study.)
Tidyman, W. F., and Brown, Helen. "The Extent and Meaning of the Loss in Transfer in Column and Dictation Spelling," *Elementary School Journal,* XVIII, (November, 1917), 210-215.
Hunkins, R. V. "An Experiment in Column vs. Dictation Spelling," *Elementary School Journal,* XIX, (May, 1919), 689-699.

[2] Winch, W. H. "Additional Researches on Learning to Spell," *Journal of Educational Psychology,* VII, (February, 1916), 93-110.

[3] Hawley, W. E., and Gallup, Jackson. "The 'List' vs. the 'Sentence' Method of Teaching Spelling," *Journal of Educational Research,* V, (April, 1922), 306-310.

[4] McKee, Paul. "Teaching Spelling by Column and Context Forms," *Journal of Educational Research,* XV, (April and May, 1927), pp. 246-255 and 339-348.

The Problem.—The present study is concerned with another aspect of "column vs. context" spelling not dealt with in the foregoing experiments. Do pupils know the meanings of words tested in context better than they do those tested in column form? The study here reported was carried out in order to answer this question.

Method.—The method of experimentation consisted of testing pupils with a list of words in column form and an equivalent list in context form. Methods of presentation and study were the same for both forms. In order that no advantage might inhere in one form due to priority of testing, the pupils were divided into four groups and the order of presentation was varied for each group. The method of rotation is set forth in a later section of the report.

Selection of Words.—In order to investigate the efficiency of the two methods of testing stated in the above question, it was necessary to have tests in which the words were balanced not only as to difficulty of spelling, but also as to difficulty of meaning. Consequently, two preliminary tests of the multiple-choice type were constructed and given to over seven hundred high- and low-5 pupils. The difficulty of the 214 items comprising the two tests was determined by computing the percentage of pupils passing each item correctly. By utilizing the percentages of correct spelling given in the Iowa Spelling Scales for Grade 5, it was possible to construct two tests of equal spelling and vocabulary difficulty.[5] Each test consisted of forty items. Two equivalent items, one for each test, are given below:

notify (to say, act of observing, to inform, to notice)
usual (practical, valuable, using, customary)

The means and standard deviations of each test are given in Table I.

TABLE I

MEANS AND STANDARD DEVIATIONS OF VOCABULARY TESTS AB AND CD ON SPELLING AND VOCABULARY DIFFICULTY

	Spelling Difficulty*		Vocabulary Difficulty	
	AB	CD	AB	CD
Mean	65.0	64.8	42.8	42.7
S.D.	9.1	9.0	14.5	14.6

* Based on percentages given in Iowa Spelling Scales.

The correlations of tests AB and CD are given in Table II. Since the present investigation deals with group measurement, the correlations here reported seem adequately high.

[5] Ashbaugh, Ernest J. *The Iowa Spelling Scales.* Bloomington, Ill.: Public School Publishing Company, 1922.

TABLE II

CORRELATIONS OF TEST AB WITH TEST CD

	r	P.E.r
383 High- and Low-5 Pupils	.742	.015
214 High-5 Pupils	.736	.018
169 Low-5 Pupils	.769	.021

Procedure.—Since tests AB and CD were to constitute the spelling words taught over a period of several weeks, it was deemed advisable to break up each test into two parts thus making four tests of twenty items each. The four tests are designated as A, B, C, and D. The eighty words in these tests were taught over a period of four weeks to eight high-5 and eight low-5 classes in several of the Atlanta and Houston Public Schools. The column form of testing was used in tests A and B. The words in C and D were tested by sentence dictation. Each dictation sentence employed the particular meaning of the word contained in the vocabulary test. The dictation sentences were also used when presenting the spelling words for a given week. This was done with the words in test AB as well as with those in test CD.

The order of presenting the words in each of the four tests is set forth in Table III. This table shows that the sixteen classes were divided into four groups. Each group consisted of two high-5 and two low-5 classes. The classes in Group I were taught the words in test A the first week, test B the second week, test C the third week, and test D the fourth week. The order of presentation for Groups II, III, and IV may be read in like manner. The rotation of the tests was an attempt to eliminate the time factor. The numbers at the foot of the table indicate how many pupils were present for all tests.

TABLE III

ORDER OF PRESENTATION OF WORDS IN VOCABULARY TESTS A, B, C, AND D

	Group I	Group II	Group III	Group IV
First Week	A	B	C	D
Second Week	B	A	D	C
Third Week	C	D	A	B
Fourth Week	D	C	B	A
No. High-5 Pupils	47	57	55	55
No. Low-5 Pupils	41	42	35	51

The method of conducting the experiment will be made clearer by examining the directions for one group. The only difference in the directions for the four groups is in the order of presenting the words as was pointed out in Table III. It will be noticed in the directions given below for Group I that provision was made for giving each vocabulary test a second time one month later. Thus it was possible to compare the results

on vocabulary tests A and B with those on C and D for both immediate and delayed recall.

GROUP I

Directions for Conducting Experiment

The directions below are given in the order in which they are to be followed:

First Week:

Monday

1. Present the *underscored* words in Spelling Test A. (See directions under, "Teaching Procedure for Column-Dictation Spelling.")
2. Follow presentation of underscored words in Spelling Test A immediately with test on words. (See directions under "Teaching Procedure for Column-Dictation Spelling.")

Tuesday, Wednesday, Thursday

Follow directions given under, "Teaching Procedure for Column-Dictation Spelling."

Friday

1. First, give Spelling Test A.
2. Follow immediately with Vocabulary Test A.

Second Week:

The experimental procedure for the second week is the same as the procedure for the first week, except that Spelling Test B and Vocabulary Test B are now used. Be sure that Spelling Test B precedes Vocabulary Test B on Friday. Be sure, also, in spelling tests on Monday, Wednesday, and Friday, to dictate *only* the *underscored* words.

Third Week:

Monday

1. Present *underscored* words in Spelling Test C. (See directions under, "Teaching Procedure for Sentence-Dictation Spelling.")
2. Follow presentation of words in Spelling Test C with Sentence dictation test. Have pupils write each dictated sentence. (See directions under, "Teaching Procedure for Sentence-Dictation Spelling.")

Tuesday, Wednesday, Thursday

Follow directions given under, "Teaching Procedure for Sentence-Dictation Spelling."

Friday

1. First, give Spelling Test C. Have pupils write each dictated sentence.
2. Follow immediately with Vocabulary Test C.

Fourth Week:

The experimental procedure for the fourth week is the same as the procedure for the third week except that Spelling Test D and Vocabulary Test D are now used. Be sure that Spelling Test D precedes Vocabulary Test D on Friday. Be sure, also, in spelling tests on Monday, Wednesday, and Friday to have pupils write the sentences dictated by the teacher.

Fifth Week:
Friday
Give Vocabulary Test A.
Sixth Week:
Friday
Give Vocabulary Test B.
Seventh Week:
Friday
Give Vocabulary Test C.
Eighth Week:
Friday
Give Vocabulary Test D.

Besides the directions for conducting the experiment each teacher was provided with spelling directions for teaching and testing the words. There were two sets of teaching directions, one for column dictation and one for sentence dictation. The column-dictation directions were to be used with the words in tests A and B and the sentence-dictation directions with the words in tests C and D. The only difference in the two sets of directions was in the method of testing. The A and B words were tested in column form, the C and D words in sentence form. The column-dictation directions follow:

TEACHING PROCEDURE FOR COLUMN-DICTATION SPELLING

Monday

1. Write on the board the underscored word in the first sentence.
2. Pronounce it distinctly by syllables.
3. Have the children pronounce it in the same way.
4. Read the sentence containing the word to the class in order that they may sense the meaning of the word.
5. Repeat the above steps (steps 1, 2, 3, and 4) with each of the succeeding underscored words. Keep the list of words on the board for the remainder of the week. Cover the words with a wall map or other means during the tests on Monday, Wednesday, and Friday.
6. Test the pupils on the twenty *underscored* words. Dictate each of these words. Have the pupils write them in column form. Do *not* dictate the the sentences.
7. Allow children to correct their papers and to list misspelled words for study on Tuesday.

Tuesday

1. Excuse from study pupils who spelled all words correctly on Monday.
2. Direct each pupil to study the words *he missed* on Monday.
3. The following procedure for individual study should be used:
   a. Look carefully at the word pronouncing it silently by syllables.
   b. Close your eyes and try to visualize the word.

c. Write the word and compare it with the correct spelling on the board. Do this until you have written the word correctly three times.
d. Re-study the words you found most difficult.

Wednesday

1. Re-test *entire* class on all words as on Monday.
2. Allow children to correct their papers and to list misspelled words for study on Thursday.
3. Direct children to note progress from Monday to Wednesday.

Thursday

1. All pupils who spelled all words correctly on Wednesday are excused from study.
2. Use same procedure in studying as was used on Tuesday.

Friday

1. Follow procedure used on Wednesday.
2. Direct pupils to note progress during the week.

In addition to the above materials, each teacher was given four sets of dictation sentences to be used in the presentation of the words. The dictation sentences containing the words in tests C and D were also used for testing the spelling of those words. The sentences containing the words in test A are given below:

SPELLING TEST A

1. He *expressed* his ideas clearly.
2. Please *notify* the clerk that we are here.
3. It is a *worthy* act to help the needy.
4. They *limited* his allowance.
5. The *surface* of the wood was highly polished.
6. The goods were sold at a *reduction*.
7. The *average* person is not wealthy.
8. Your conduct will *entitle* you to a reward.
9. An *investment* in oil is risky.
10. Unless you make an *attempt* you cannot do it.
11. The pupil was *confident* that he acted wisely.
12. The *base* of the statue was strong enough to support it.
13. The man was required to *inspect* each article carefully.
14. He was *inclosing* the yard with a wire fence.
15. The doctor will *require* him to sleep ten hours.
16. He *procured* enough worms to go fishing.
17. His argument *convinced* us that it was right to go.
18. They will *confer* about the peace plans.
19. The apartment lease *expired* last month.
20. I *prepaid* the postage on your package.

Results.—The comparisons to be presented are between the results on vocabulary test AB and CD. It will be remembered that tests AB and

CD were broken up for testing purposes into A and B, and C and D, respectively. The comparisons given in the following table are for all high- and low-5 pupils who took the eight tests.

TABLE IV

DIFFERENCE BETWEEN THE MEAN VOCABULARY SCORES MADE BY 383 HIGH- AND LOW-5 PUPILS WITH THE COLUMN METHOD AND THOSE MADE WITH THE SENTENCE-DICTATION METHOD

	First Testing			Second Testing		
	Column AB	Sentence CD	Difference	Column AB	Sentence CD	Difference
Mean	21.4	24.2	2.8	23.8	26.2	2.4
S.D.	7.0	5.8	.242	7.0	6.0	.229

Table IV shows a mean of 21.4 on vocabulary test AB as compared with the mean 24.2 obtained on test CD. The difference between the means is 2.8 in favor of the sentence-dictation method. The standard deviation of the difference is .242. The difference between the means of the two tests when given a month later is 2.4. Since both differences are more than ten times their sigmas, it would appear that the method of testing spelling by sentence-dictation is superior to the column-dictation method from the standpoint of the ability of pupils to recognize one meaning of a word.

Tables V and VI give the results for 214 high-5 and 169 low-5 pupils, respectively. All of the differences are in favor of the sentence-dictation method and exceed their sigmas by four or more times.

TABLE V

DIFFERENCE BETWEEN THE MEAN VOCABULARY SCORES MADE BY 214 HIGH-5 PUPILS WITH THE COLUMN METHOD AND THOSE MADE WITH THE SENTENCE-DICTATION METHOD

	First Testing			Second Testing		
	Column AB	Sentence CD	Difference	Column AB	Sentence CD	Difference
Mean	22.9	25.3	2.4	25.7	27.1	1.4
S.D.	6.2	5.2	.531	6.4	5.5	.280

TABLE VI

DIFFERENCE BETWEEN THE MEAN VOCABULARY SCORES MADE BY 169 LOW-5 PUPILS WITH THE COLUMN METHOD AND THOSE MADE WITH THE SENTENCE-DICTATION METHOD

	First Testing			Second Testing		
	Column AB	Sentence CD	Difference	Column AB	Sentence CD	Difference
Mean	19.4	22.7	3.3	21.4	25.0	3.6
S.D.	7.5	6.3	.369	6.9	6.4	.377

Since many of the words used in the experiment are too difficult for a fifth-grade spelling vocabulary, it may be objected that different results

would have been obtained with an easier list of words. Assuming that difficult spelling words are also difficult from the standpoint of meaning, the writers purposely chose words that would be sufficiently discriminating to bring about a rather wide range of ability. A further objection may be made that pupils already know the meanings of words they learn to spell because the words have in large part been chosen from the written vocabularies of children. This objection may be met by the findings of a recent study reported by Vogel and Washburne.[6] These writers describe a method for determining which spelling words pupils do not know the meanings of and conclude "that the meanings of many words are unfamiliar to the children who are learning to spell them." It seems possible that an appreciable difference between the two methods would result if the experiment were repeated with the actual spelling words for a given grade.

Conclusion.—The present study shows the superiority of sentence-dictation spelling over column-dictation spelling in teaching the meanings of words. This investigation does not show that the pupils were more able to recognize other meanings of the words or to use the words in sentences of their own construction. These are problems for further investigation.

[6] Vogel, Mabel, and Washburne, Carleton. "A Year of Winnetka Research," *Journal of Educational Research*, XVII, (February, 1928), 90-101.

COMPARATIVE SCHOOL MARKS OF SEPTEMBER AND JANUARY JUNIOR HIGH-SCHOOL ENTRANTS

ORLIE M. CLEM
Syracuse University
and
J. K. WALKLET
Franklin, New York

Do pupils who enter the junior high school in September do better work than those who enter in January, or vice versa? There seems to have been current a rather vague and ill-defined, yet real, opinion among teachers and administrators that January high-school entrants are somewhat irregular, if not abnormal. The investigation here summarized was conducted in one school, primarily to determine whether there is any scientific basis for such an assumption. Secondarily, the investigation was made to show the relationship between regents' marks and teachers' marks for the ninth-grade pupils of a junior high school.

The investigation was conducted in the Lincoln Junior High School of Syracuse, New York. This school of one-thousand pupils seemed well adapted to the experiment because of the varied character of its pupils; it is distinctly not a class school, but seems to represent a fair cross-section of American school life. One hundred eighty-eight pupils for each group of entrants were used as a basis of comparison. The period during which these enrollments were made extended from September, 1921, through September, 1926. These one hundred eighty-eight pupils of each group represented a sampling both in years and in other selective factors, within the ninth year, for the five-year period.

The marks of the pupils were compared in English, Latin, biology, and algebra. The average of the two quarter marks was used, as term marks were not given in this school. The data for the comparisons of teachers' marks were derived from 2,439 averages and 1,718 examination marks. Inasmuch as each average represents two quarter marks, the total number of school marks, which were actually considered, was 6,596, a sufficient number to give significance to the results obtained. For classification of the marks, an interval of five, beginning at 51 percent and ending with 100 percent, was used. Those pupils falling below 50 percent were classified together. The initial comparison in each subject was made between the first-term work of the two groups; the same procedure was used for the work of the second term. Next, in order to determine

the character of the work done within the group, the two classes of entrants were compared *inter se*.

SUBJECT BY SUBJECT COMPARISONS

English.—In first-term English, the work done by one group very closely conformed to that done by the other as the medians—January, 79.9, and September, 80.5—given in Table I show. Furthermore, when the marks were grouped in intervals of five, and the percents of pupils within the interval for each class of entrants were compared, it was obvious that the relation was close. In the second term, again, but slight difference appeared between the medians. While the percents within each interval were not so close as in the first term, no sharp demarcations existed which negate the prevailing uniformity.

TABLE I

PERCENTS OF JANUARY AND SEPTEMBER ENTRANTS WHO RECEIVED VARIOUS MARKS

	Teachers' Marks											
	0–50	51–55	56–60	61–65	66–70	71–75	76–80	81–85	86–90	91–95	96–100	Median
ENGLISH:												
First term—												
January entrants	0.0	0.0	0.0	0.0	5.4	20.3	32.2	22.2	12.8	6.4	0.0	79.9
September entrants	0.0	0.5	0.5	1.1	3.2	18.2	33.1	21.4	15.0	7.0	0.0	80.5
Second term—												
January entrants	0.0	0.0	0.0	2.1	4.1	15.8	32.8	25.1	12.3	7.5	0.0	80.3
September entrants	0.0	0.0	0.0	0.7	6.9	19.2	25.2	24.9	17.3	6.8	0.0	79.2
LATIN:												
First term—												
January entrants	0.0	0.7	1.4	2.1	4.2	15.4	16.2	18.2	25.2	10.5	6.3	83.8
September entrants	2.8	0.7	3.5	4.9	8.4	10.5	18.9	16.1	18.2	11.9	4.2	81.5
Second term—												
January entrants	1.0	0.0	1.0	1.0	6.1	12.1	20.2	32.3	16.2	7.1	3.0	82.3
September entrants	0.0	0.0	0.0	2.1	3.2	12.7	18.1	21.3	26.6	14.9	1.1	83.3
BIOLOGY:												
First term—												
January entrants	0.0	1.1	2.7	5.3	10.6	25.5	19.1	19.2	11.4	5.3	0.0	77.4
September entrants	0.5	2.7	3.2	5.9	12.8	26.2	12.6	21.4	10.1	4.7	0.0	75.8
Second term—												
January entrants	0.0	0.0	0.7	2.9	5.0	30.0	32.2	20.8	8.6	0.0	0.0	77.9
September entrants	0.0	0.0	0.7	2.2	8.8	33.1	21.4	14.7	15.4	3.7	0.0	77.2
ALGEBRA:												
First term—												
January entrants	2.7	0.5	2.1	3.2	5.3	10.7	12.3	19.2	24.6	13.9	5.3	84.1
September entrants	2.7	3.2	2.7	6.4	6.4	11.7	16.5	15.5	17.6	10.6	6.9	84.9
Second term—												
January entrants	0.7	1.5	1.5	2.1	4.4	16.3	18.5	22.2	17.8	8.9	5.9	82.1
September entrants	0.0	0.0	0.8	0.8	10.6	11.3	18.2	13.7	20.4	13.6	10.6	82.7
FIRST-TERM AGGREGATE:												
January entrants	0.7	0.6	1.6	2.7	6.5	18.2	20.1	19.8	18.0	8.9	2.7	80.9
September entrants	1.4	1.9	2.4	4.6	7.7	17.0	20.3	18.7	15.0	8.4	2.7	79.7
SECOND-TERM AGGREGATE:												
January entrants	0.4	0.4	0.8	2.1	4.8	19.0	26.6	24.8	13.3	5.8	2.1	80.3
September entrants	0.0	0.0	0.4	1.4	7.7	19.6	21.1	18.3	19.3	9.3	2.9	80.9

To determine whether one group showed a tendency toward faster progress than the other, the work of the first term was compared with that of the second term. In both cases a high degree of similarity prevailed. From these data we can conclude that there was little, if any, difference in the quality of the work which the two groups did in English. Neither group made greater progress from term to term than the other.

Latin.—Only about 75 percent in either class of the entrants elected Latin, whereas practically all students were enrolled for the first-term work in the other subjects. Domestic economy or manual training was substituted in almost every case where Latin was not elected. The pupils who were not enrolled in Latin classes did not receive high marks in the other three subjects. The medians for the first term—January, 83.8, and September, 81.5—were very similar. The relatively high level probably resulted from the selective process mentioned above. When the two classes of entrants were compared, interval by interval, they were found to diverge from the steady uniformity exhibited by the corresponding comparison for first-term English. There was, nevertheless, an unmistakable drift toward uniformity.

In second-term Latin a further falling off was in evidence; only about 55 percent of the entrants elected it. In spite of this high degree of selection, the medians maintained approximately the same level as those for the first term, 82.3 and 83.3. Likewise, when the corresponding percentage of marks in the various intervals were compared, some irregularity was noticeable although a general drift toward sameness again appeared.

When the two classes of entrants were compared *inter se* term by term, it was found that in progress neither group showed any marked superiority over the other. Thus, while Latin did not disclose the general trend toward uniformity as strikingly as English, a general similarity was maintained.

Biology.—This is a subject which differs from the languages both in type and in the fact that it is a regents' subject. With one exception, all of the pupils took biology. Table I shows that while the medians resemble each other rather closely, they are appreciably lower than those for the first term in either English or Latin. One would not believe that biology is more difficult than English or Latin, so this tendency toward lower marks must be due to some other influence. Data, which will be discussed later, show that the general level of the marks in

regents' examinations is much lower than that of the school marks. It is therefore probable that the tendency to mark more strictly in this subject is due to an effort to make the school marks approximate the regents' marks.

While the marks of the two classes of entrants did not correspond as closely as the English marks, a closer degree of similarity existed in the first term of biology than in the case of Latin. The medians were very close, as reference to Table I shows. For the second term they maintained a comparatively low level, but they corresponded more closely than those of the first term. Since biology probably does not grow increasingly difficult week by week, the greater number of mediocre marks apparently indicates that biology is marked even more rigidly in the second term. This fact, taken in connection with the finding that most of the failures occur in the first term, points to a conscious effort of grooming for the regents' examinations. Biology evidently shows less tendency toward uniformity than English, but a greater tendency than Latin.

Algebra.—Algebra, the last subject considered is a regents' subject also, and practically all of the pupils were enrolled for the first term. In the second term, the usual falling off takes place, but it is somewhat greater than in any of the subjects previously examined. The data differ from that previously considered in their exceptionally wide range of distribution. Every grade interval is represented in both terms, a range not shown in any of the preceding cases.

Algebra, in all of the comparisons, revealed some rather wide discrepancies which, however, were no greater than those noted for Latin. Allowing these divergences full weight, there still remains a pronounced tendency toward similarity.

SUMMARY OF THE COMPARISONS

In the preceding analysis of each separate subject any consistent tendency toward the superiority of one class of entrants over the other is found to be conspicuously lacking. A general drift does seem apparent—however, indistinct at times—toward a uniform standard of attainment for both classes of pupils. The evidence of this is so slender, in some instances, that it would be unsafe to make a generalization which would apply without modification to each individual study. Observing, then, that no definite prediction could be made for any given subject and that one group was just as apt to precede as to follow by a small number

of points, it seemed advisable to find the relation between the work done by the two groups as a whole. This was done by comparing the total number of marks in each corresponding group.

A marked similarity exists between the two groups in the first term, so marked, in fact, that when the data are graphed there is presented what is virtually a double normal curve. In the second term, the coherence is not very marked, but the divarication is not so great as to impair the general effect of similarity. The *inter se* comparisons show no wide variations, but it was noted that in all of the subjects taken together the January entrants do their best work in the first term, while the September entrants do theirs in the second. This is probably due to the fact that these terms follow immediately after a previous semester of school work, while the second term of the January entrants and the first term of the September entrants follow the long summer vacation. The difference is slight, however, thus indicating that vacation does not seriously disturb the educational continuity.

When the work of the entire year for both groups was compared, the resemblance was found to be very close (see Figure 1). For these groups

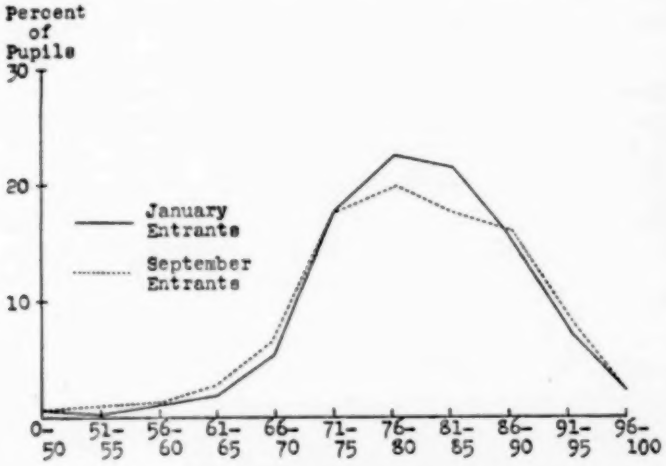


FIGURE 1. AGGREGATE MARKS FOR THE ENTIRE YEAR

of pupils in this school the following conclusions seem justifiable in terms of this study: In general, there is no essential difference in the scholarship of the January and September ninth-grade entrants. Neither group shows any marked superiority over the other either in scholarship

or progress. Time of entrance is not a governing factor in ninth-grade school work.

COMPARISON OF ENTRANTS ACCORDING TO REGENTS' MARKS

The material in this part of the investigation involved a new factor, namely, regents' marks. The entrants were first compared on the basis of the regents' marks. A general comparison was then made between the work done by pupils exempt from the final examinations and the "non-exempts."

Since biology and algebra are the only regents' subjects taken in the ninth grade, the first phase of the investigation is necessarily confined to these studies. It was found that in biology the two groups averaged about the same. In algebra there was a similarity, but when the data were graphed the resemblance to a normal curve was slight. The middle of the curve was depressed instead of being elevated with the passing mark as the peak. This seems to cast some doubt upon the old type of algebra examination as a precise measure of varying ability. According to these results, it tended to differentiate, merely, between the good and the poor but was not a scale.

The regents' marks in both algebra and biology consistently maintained a lower level than the school marks. When the data were graphed, the contour of the curve for biology showed a marked resemblance to a normal curve, thus justifying the conclusion that the regents' examinations in biology seem to be good scales for grading the abilities of the pupils. The findings were just the opposite for algebra. Here the general average curve more nearly approached the normal, while the regents' curve was very erratic. This substantiates the conclusion drawn in the preceding paragraph.

Regents' marks were compared with the marks obtained in the examination at the end of the first term. Although each regents' mark was compared with its own corresponding term-examination mark, all of the comparisons did not involve the same year. Any general drift that was found to result would, therefore, be significant. The number of cases considered decreased, however, by the fact that not all pupils took the term examinations. Some were exempt. The findings which follow, then, do not pertain to the superior pupils.

Of the "ordinary" pupils in biology, two out of three of both classes of entrants made lower marks in the regents' examination. This tendency was even more striking in algebra; four out of five of the January

entrants were graded lower in the regents' examination. The September entrants show an even greater decrease. The results in this latter case are somewhat influenced by the examination of June, 1925, which, because of excessive failures, was later invalidated by the State Department. Although the data do not warrant any sweeping conclusions, it is evident that non-exempt pupils tend to get lower marks in the regents' than in the school examinations.

While the foregoing facts deal with those pupils who have to take both examinations, it does not embrace that class who are exempt from the term examinations and take only the regents'. It would be futile to speculate what these pupils might do if they had taken the term examinations. Exemption assumes a superiority, however, and this assumption should be checked. To do this, the regents' marks of the exempt pupils in each group were compared with those of the non-exempts of the same group. This phase of the investigation showed that the exempt pupils do appreciably better in regents' than the non-exempts. In other words, the results show that the teachers' judgment is confirmed by the regents' marks, and the not uncommon statement that "the best pupils very frequently fail on account of the unfairness of the regents' examinations" is refuted.

In a further investigation of the subject of exemptions, the number of exemptions obtained by the January and September entrants were compared in each subject. On this comparison, the following conclusions were based: (1) Neither class of entrants, as a whole, consistently leads the other in exemptions. (2) In English and Latin—subjects in which exemptions may be earned in either term—the January entrants lead in the first term, while the September entrants exceed in the second term. The *inter se* comparisons show a like alternation. This is in accord with a tendency which has been previously noted and discussed. (3) Latin shows a greater average of exemptions than does English. Due to excessive elimination, however, the group taking Latin is more or less selected. (4) Of the two regents' subjects in which only one-term exemptions are possible, algebra leads by a considerable margin. This would be expected, for it has been shown that in this subject a sharp division exists between the good and the poor pupils.

ON THE TECHNIQUE OF ESTABLISHING THE "GOODNESS" OF AN INTELLIGENCE TEST

JACK W. DUNLAP
Stanford University

THE following discussion on the technique of certain methods of evaluating tests is the result of examining somewhat critically the methods proposed by Dr. V. A. Jones in his recent monograph, *The Effect of Age and Experience on Tests of Intelligence.*[1]

The purpose of Dr. Jones' monograph is to investigate the influence of chronological age upon mental age. The methods used are those of partial correlation and the K index. The latter will be explained presently.

It is, of course, important to measure mental maturity as distinct as possible from the effects of chronological age, environmental factors, health, etc. Jones has a group of twenty-six tests which he ranks from one to twenty-six as he considers best for measuring MA independently from CA.

The total population used in this study is 487. The author classifies the population into three groups: (1) Richmond Sampling, 208 white children from the lower half of the fifth grade. (2) Miscellaneous Sampling, grades 5-7, population 179. (3) Miscellaneous Sampling, grades 3-4, population of 100.

The tests studied were the subtests, totaling twenty-six, from the following: National Intelligence Test, Scale A; National Intelligence Test, Scale B; Haggerty Intelligence Examination, Delta 2; Otis Group Intelligence Scale, Advanced; and the Otis Group Intelligence Scale, Primary, for grades 3 and 4.[2]

Jones uses two criteria to establish the goodness of a test; first, that the partial correlation between $r_{s\ ma.ca}$ should approach 1, where s can stand for any subtest; second, that the partial correlation $r_{s\ ca.ma}$ should approach zero.

---

[1] Contributions to Education, No. 203. New York: Teachers College, Columbia University.

[2] The author made a study of Jones' methods of sampling, criticizing several points of procedure in the earlier work. In choosing the MA of the child from the four tests, however, he chose the median age, since it is the easiest. It is easy, but it is not the best method for arriving at the real MA.

He uses the alienation coefficient of these two partials in ranking each subtest by his K index. This index is furnished by the two formulas:

(a) $k_{s\ ma.ca} = \sqrt{1 - r^2_{s\ ma.ca}}$
The nearer this value is to 1, the poorer.

(b) $k_{s\ ca.ma} = \sqrt{1 - r^2_{s\ ca.ma}}$
The nearer this value is to 0, the poorer.

Another way of representing the latter formula is to say that the greater the value $1 - \sqrt{1 - r^2_{s\ ca.ma}}$ the poorer the test. The sum of the measures furnished by formulas (a) and (b) in the latter form gives the K index.

The larger the K index the poorer the test. A slight correction was made because of the fact that every test is used in calculating MA and thus in correlating $r_{s.ma}$ we tend to get a spuriously high r. The author's conclusions show that he considers these criteria the most important product of the thesis. He has gone to a great deal of trouble in calculating them and even to much more trouble to get the K index.

Believing for practical work that if the first condition is fulfilled the second will tend to be approximated very closely, the writer ranked the tests by means of $r_{s\ ma}$ getting almost identical results with the ranking obtained by Jones using the K index. It can be shown by a few simple steps in algebra that if $r_{s\ ma} = 1$, then of necessity the partial $r_{s\ ma.ca} = 1$.

$$r_{s\ ma} = 1$$

Therefore $r_{a\ s} = r_{a\ ma}$

Where a means (anything) a variable to be used in both r's.

$$\text{Then } r_{s\ ma.ca} = \frac{r_{s\ ma} - r_{s\ ca} r_{ma\ ca}}{\sqrt{1 - r^2_{s\ ca}}\ \sqrt{1 - r^2_{ma\ ca}}} = \frac{1 - r^2_{s\ ca}}{1 - r^2_{s\ ca}} = 1$$

One may infer from this that if our correlations between subtests and MA are high we have the most important criterion, the other being so small in its effects as to be negligible for all practical purposes. Accordingly, the most important thing is to get high r's between s and MA.

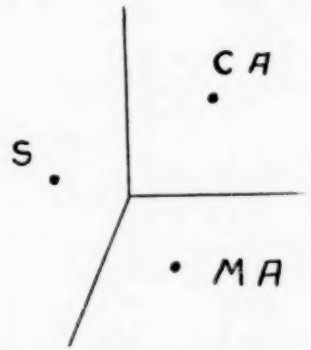
If such a condition holds true even approximately we should be able to rank the tests accurately by means of the r_{sma}, corrected for the amount of error introduced by using the test against an MA which was in part derived from the test. The correction used by the writer in arriving at his ranking of the r's was the square of the coefficient of correlation found by Jones to be due to the fact that the subtest forms a part of MA. This varies for each group of tests so it cannot be treated as a constant in the ranking. In order to make his own ranking com-

parable with that of Jones, the writer weighted each group as Jones did, then divided as Jones indicated by the appropriate constant.

The five highest tests found by Jones' complex method are identical with the five found by the writer and they are in the same order. The five lowest found by Jones are the same as found by the writer except that the tests ranking twenty-four and twenty-five have been reversed. Calculating rho between the two rankings we get .993, which expressed in terms of r is .994. In view of the above and considering that one need calculate only the $r_{s\,ma}$, eliminating the necessity of getting $r_{s\,ca}$, $r_{ma\,ca}$, and the partials, it seems that the method proposed by Jones is unnecessary technique. If all the r's, however, are close to zero and if one wanted the best method of ranking a group of tests, regardless of labor, Jones' technique would be the proper one.

The question immediately arises whether it is a matter of chance that the rankings are so nearly identical or whether there is a mathematical relationship that holds constant throughout. If there is such a relationship, under certain conditions we do not need to calculate either of the partials, but need merely correct and weight the $r's_{ma}$ by the best method indicated by the data. That there is such a relationship between his variables and that the similarity of rankings is not due to chance will now be shown.

Let us think of our three variables with population constant as here, and remember that since we are using the same population in all our intercorrelations, MA and CA do not change. The correlations of the subtests with MA and CA do vary.[3]



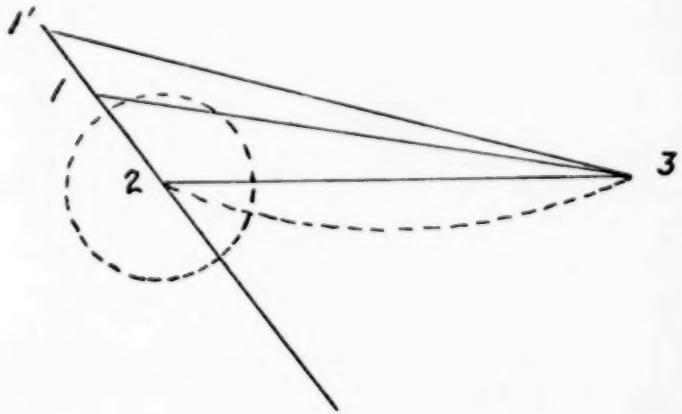
[3] The writer is indebted to Dr. Harold Hoteling, for assistance with the following development.

We can think of our cases for a subtest plotted in n dimensional space, and all concentrated at a point, say s, in a plane. The same can be thought of for each of the other two variables, giving us three points in three planes.

By weighting and evaluating we can so shift the mean that we will place it at zero for all three. Now let us conceive of a sphere with this as the center, and pass lines through the points in the planes, and consider the curved surface made by connecting the points at which they cut the surface of the sphere.

The cosine of the ∠s ca ma equals $r_{13.2}$. From this point s, MA, and CA will be represented by 1, 2, and 3 respectively. The cosines of the angles of the spherical triangle represent the partial correlation coefficient. The shorter the arc between two points the greater r_{xy} between the two variables represented by the two points. Now r_{23} is fixed in this case at approximately —.21.

We wish to show that the tests rating low in numerical value by the criteria $1 - K_{13.2} + K_{12.3}$, are the same as the tests rating high by the



criteria r_{12}. Now $1 - K_{13.2} + K_{12.3}$ can be written $K_{12.3} - K_{13.2}$, by discarding the 1 as a constant, and it is seen that the tests ranking high in numerical value are the same as the ones ranking high in r_{12}. In order to be consistent we will use K_{12} instead of r_{12}, and now a test ranked high by the criteria $K_{12.3} - K_{13.2}$ should be low by K_{12}. That is, the two vary directly together when compared in terms of r_{12} and the value $K_{12.3} - K_{13.2}$. This is shown by the diagram on page 370.

Straight lines are used to represent the arc because of ease of illustrating. The arc 2 3 is fixed. Arc 1 2 is a variable and may lie anywhere on the dotted circle about 2. Angle 1 2 3 is fixed, but as the arc moves out from 1 to 1′ it can easily be seen that when angle 1 3 2 increases angle 2 1 3 decreases and vice versa. Let us consider two examples given from the data in Jones' monograph.

(a)	(b)
$1 = s$	$1 = s$
$2 = MA$	$2 = MA$
$3 = CA$	$3 = CA$
$r_{12} = .26$	$r_{12} = .72$
$r_{13} = -.13$	$r_{13} = -.188$
$r_{12.3} = .239$	$r_{12.3} = -.188$
$r_{13.2} = -.067$	$r_{13.2} = -.708$
$\angle 123 = 93° 51'$	$\angle 123 = 92° 52'$
$\angle 132 = 76° 12'$	$\angle 132 = 45°$
Arc $12 = 75°$	Arc $12 = 44°$

$r_{23} = -.21$ is the same for both.

Now we get this picture:

	$K_{12.3} - K_{13.2}$	k_{12}
(a)	$.996 - .973 = .023$	.965
(b)	$.998 - .707 = .291$	.695

This is exactly what we would expect from a trigonometric study of correlation, that the K coefficients would vary inversely with $r_{s\,ma}$, and the author's criterion is a function of $r_{s\,ma}$.

It is perfectly conceivable, however, that by chance when rating a group of tests, the r's might form some sort of a spiral around 2 in which case the criteria r_{12} would hold only approximately and then only for the outlying cases. The higher r_{12} the more accurate the rankings, and when the r's are above .35 it is a reliable method. This last is an empirical finding and as such may be subject to some error.

The main contribution of the study in question is in the need for a better method of expressing mental maturity and of measuring it. The author also shows statistically that at certain levels a test is better than at others. It is hoped that in the future other workers will go into this field and give us more and better tools.

EDITORIALS

JOHN DEWEY'S SEVENTIETH ANNIVERSARY

It is quite fitting that a magazine bearing the title *Journal of Educational Research* should take note of the recent celebration of the seventieth birthday of John Dewey. Few philosophers have influenced so directly and so immediately the attitude of teachers and parents toward children.

Perhaps it is indicative of a new turn in American culture that representatives from virtually every walk of life should meet in New York on the occasion of John Dewey's birthday and review in his presence the influence of his thought upon contemporary life. On Friday evening, October 18, educators from all parts of the United States and abroad gathered together to survey his contributions to education. At this meeting both Dr. I. F. Kandel and Dr. Jesse H. Newlon spoke. Dr. Kandel emphasized Dewey's influence upon education abroad, while Dr. Newlon stressed "John Dewey's Influence Upon the Schools" in this country.

"Dewey," said Dr. Newlon, "found the American school strongly intrenched in custom and tradition, still in the grip of a treadmill pedagogy that at best frustrated the child's normal impulses at almost every point —a formal, static and unimaginative institution." Dr. Newlon then proceeded to show how Dewey in focussing attention upon certain essential factors in the theory and practice of education has given direction to a revolution in educational aims and procedures. Dewey's books, *School and Society, The Child and the Curriculum, Interest and Effort in Education, How We Think, Moral Principles in Education, Democracy and Education,* constitute the handbooks of this revolution in education. They have given validity to the principles that education must focus attention upon the nature and the needs of the child and thus substitute child growth for subject matter as a central factor in school activities; that education is a process of experiencing, and in Dewey's words "Learning is active. It involves organic assimilation starting from within;" that genuine interest involves both "interest" and effort because it recognizes "the identity of the fact to be learned or the action proposed with the growing self;" and, finally, that the school is a community whose processes are social processes, not different from the social processes that go on outside of school.

Dr. Newlon then proceeded to trace the transforming character of these principles upon the curriculum, upon method, pupil activities within schools, moral education, religious education, building and equipment, the spirit of administration and supervision, and upon the relations between the school on the one hand and the home and other social agencies on the other.

On Saturday morning, October 19, consideration was given to the far reaching consequences of Dewey's philosophy and at one o'clock on the same day some two thousand men and women sat down to a luncheon held in his honor at the Hotel Astor. The guest list on this occasion contained the names of many of our most notable leaders in education, philosophy, social work, political economy, law, the press, etc. As one glanced from the formidable array of distinguished faces at the speakers' table to the modest and retiring, if not embarrassed, figure of John Dewey at its center, he sensed the appropriateness of the morning editorial in the *New York Herald Tribune*. This read in part as follows:

"It is a pleasure to take part in the celebration to-day of the seventieth birthday of John Dewey, America's distinguished philosopher. Why he is distinguished, many who recognize the fact would none-theless be unable to explain. His philosophical books puzzle at times even philosophers. His educational theory, widely influential as it is, has many to question it and many to whom it remains inspiring but ambiguous. His influence is far greater than the number of his readers, and his style is not memorable or quotable as was that of William James. What is the secret of his prestige and power?

"That answer is almost a banality. It is the sincerity and simplicity of the man that makes his philosophy so conquering and so provocative. These qualities stamp the man's manner as well as his philosophy. And in Philosophy they are very great and very rare virtues. It is his honesty that led Dewey to abandon the technical concerns of philosophy—in which by the way he is most adroit—and turn to the treatment of philosophy as a human enterprise, originating in human perplexity and helpful in the solution of human problems. It is his simplicity that taught him to look at the human mind not as an abstract entity but as a deliberate method of action of a human animal living harassed in a changing world. His whole conception of education is simply that of the inculcation of the habit of disciplined intelligence and responsible freedom in the young. His theory of democracy is that of a society of intelligently communicating and freely related human beings. His philosophy of nature is that of a growing and changing universe in which human intelligence can turn confusion into order, and nature into art. Professor Dewey has done more than any other man to transform

education from the learning of things by rote at a desk to the learning by doing of the modern school. There is no question that his doctrine of free coöperative intelligence has been the basis of any number of progressive political movements and that he has turned philosophy from the pursuit of metaphysical chimaeras to the treatment of genuine human issues."

The title of John Dewey's last book is *The Quest for Certainty*. This is no doubt prophetic. Dewey's most lasting contributions to thought will not consist of formulated statements regarding the final meaning of life. He has not ventured to paint the El Dorado of human strivings. On the contrary he would probably agree in large measure with the dramatic concluding paragraph of Robert Louis Stevenson's Essay on El Dorado:

"A strange picture we make on our way to our chimaeras, ceaselessly marching, grudging ourselves the time for rest; indefatigable, adventurous pioneers. It is true that we shall never reach the goal; it is even more probable that there is no such place; and if we lived for centuries and were endowed with the powers of a god, we should find ourselves not much nearer what we wanted at the end. O toiling hands of mortals! O unwearied feet, traveling ye know not whither! Soon, soon, it seems to you, you must come forth on some conspicuous hill-top, and but a little way further, against the setting sun, descry the spires of El Dorado. Little do ye know your own blessedness; for to travel hopefully is a better thing than to arrive, and the true success is to labor."

Instead of attempting to describe the final goal of life Dewey has concentrated upon showing men how they may construct the tools and instruments which will make real in this world human aspirations and human ideals, and which test out and validate the ingredients of the good life. His career has revealed the sincerity of his own words: "Better it is for philosophy to err in active participation in the living struggles and issues of its own age and times, than to maintain an immune monastic impeccability." His chief concern has been not to point out a final meaning of the human drama but to indicate ways in which men may discover for themselves the meaning of their experiences in the course of their living. His life has been a quest for certainty of such a character that, by virtue of the fact that he has lived, both the quest for certainty and the nature of certainty itself has taken on a significance hitherto unappreciated.

V. T. T.

REVIEWS

Address all communications relative to book reviews to Editor E. J. Ashbaugh, Miami University, Oxford, Ohio.

STILLMAN, BESSIE W. *Training Children to Study.* New York: D. C. Heath and Company, 1928. xix + 247 pp.

Three main approaches to the study problem have been common during the last several years. One is the statement of principles of how to study; a second is the development of exercises, questions, and study guides by which the student is to prepare his lessons; and the third is the development of actual teaching procedures by which the teacher can develop permanent habits of efficient studying. The book mentioned above is of the third type. It distinctly places the emphasis on learning how to study rather than merely on getting lessons learned.

The book is written for teachers of the fifth, sixth, and seventh grades, and, therefore, fills a definite need which has been more or less neglected by most of the books on study habits, which have been written for high-school and college students. The problem of teaching how to study in the elementary grades is different from that in the high school because the offering of a special course in the subject does not seem to be feasible and because more responsibility must fall upon the individual classroom teacher. If every teacher were an artist at the teaching of study habits there would be no need for regular courses in the subject, either in elementary or high school. Such a situation is not likely to be generally achieved any time in the near future, but Stillman's "Training Children to Study" is a distinct step in that direction.

A very striking feature of the work is the enormous number of actual classroom situations which are reproduced for the reader. Numerous classroom conversations are given in the exact words of the students.

Throughout the book much emphasis is given to the development of thinking. There are numerous and interesting examples of the way in which students are encouraged to gather data and defer decisions until they have the necessary foundations for them. One interesting chapter shows how the children are taught to ask questions rather than answer those asked by the teacher.

On the whole it pictures a goal or ideal so far above the level on which teaching is ordinarly done that it ought to be a very effective stimulus for teachers to approach their work from a new angle.

C. C. CRAWFORD

University of Southern California

ROBINSON, DANIEL SOMMER. *An Anthology of Recent Philosophy.* New York: Crowell Company, 1929. xii + 674 pp.

Two types of discussion in any field have an abiding interest: one, the classic works which constitute the high points in the development of the subject; the other, current opinion on its fundamental features.

For satisfaction of the former we turn to the works themselves of the leading writers or to volumes of selections from the classic thinkers. Help in meeting the

latter is offered by Dr. Robinson's book. Opinion from the time of Royce and James to the present, covering both essential problems and typical points of view, is represented in brief selections of from three to twenty-five pages. A short analysis stating the setting in which the selection occurs and summarizing the thought of the statement as a whole, precedes each passage. An appendix gives biographical and bibliographical sketches of the author.

After an introductory section entitled Orientation, the field of contemporary philosophy is divided into the views of Idealism (covering Absolute Idealism, Spiritual Monism or Creative Evolutionism, Spiritual Pluralism or Theistic Personalism), Realism (Monistic and Pluralistic), Pragmatism (Experimentalism, Instrumentalism, Humanism, Fictionalism), and a group of others not easily classifiable under these headings (Mechanism, Vitalism, Superman Philosophy, Agnosticism, and Cultural Pluralism).

The authors are for the most part from English speaking countries, there being thirty-seven Americans, nineteen British and one Hindu, four German, one French, and one Italian. The implied suggestion that philosophy is most active in America, even that perhaps little attention is being given to it in European countries, the Orient, or Spanish America, would hardly prove true, although few different movements could be added to the ones already represented. The volume represents clearly the perspective of the American observer, and gives an impression one readily gets from this angle.

The selections are well chosen, as typical specimens of the positions of the various writers, and are of a very readable sort. One who faithfully pursues his way through the volume will have added considerably to his knowledge of recent philosophy, and will have for reference many outstanding passages from the recent contributions to the literature. The book will doubtless prove very useful in courses in contemporary philosophy, aiming to give a general acquaintance with the chief current tendencies, and familiarity with the outstanding names in the field.

Ohio State University ALBERT E. AVEY

FISHBACK, ELVIN H. ***Character Education in the Junior High School.*** D. C. Heath & Company, 1928. 190 pp.

Written from the vantage point of a modern junior high school, the author has made a very readable and suggestive contribution to a rapidly increasing literature on character education. After treating briefly the modern social situation, the writer makes some reference to the characteristics of junior high-school pupils and attempts to relate these to character development through extra-curricular activities, the course of study, the personality of the teacher, etc. Finally, he has included a typical course of study for the 7th, 8th, and 9th grades.

The book is practical rather than theoretical, and has almost totally ignored the mass of modern scientific data available for guidance. Perhaps the practical values and the general readability of the book compensate to a degree for this unusual omission of really scientific material. Naturally the book finds its place for use in reading circles, and private hands, not in university courses in the field of character education.

L. A. PECHSTEIN
University of Cincinnati

WRIGHT, J. C., AND ALLEN, CHARLES R. ***Efficiency in Education.*** New York: John Wiley and Sons, Inc., 1929. xv + 405 pp.

This book appears in sixteen chapters, each dealing very specifically with some phase of efficiency in education. At the close of each chapter are found several well-put thought-provoking questions and problems. The general problems of efficiency dealt with are shown in the chapter headings among which are: Efficiency in Educational Objectives; Standards and Tests; Functioning Content; Curriculum Building; Programs; Personnel; Records and Reports; Utilization of Instructional Organizations, Building and Equipment; Procedures, Methods and Devices; Use of Coöperative Agencies; and, Securing and Maintaining Leadership.

Although a little tiresome at times, the book is rather well written. The authors approach their problem, however, too much from the vocational viewpoint. It is by no means certain that the discussions will carry over into the field of general education as readily as the authors seem to think.

A note of pessimism seems to run through the book. It doubtless is a result of contact with many deficiencies in the educational system. It is right that these shortcomings be called to mind, and yet exposing them m..y endanger the somewhat immature reader's judgment concerning the value and efficiency of the American educational system. A high degree of pessimism, as well as inadequate judgment probably, is evident in the discussion of objectives relating to ideals, attitudes, and appreciation. This discussion, however, is carried on in a wholesome way and will be stimulating to the well intentioned, sincere, and progressive administrator. While the authors point out many phases of inefficiency they are not unmindful of the fact that progress is being made and that it is exceedingly difficult in many parts of our educational program.

The book is well adapted for use as collateral reading in advanced courses in administration and supervision. It does not lend itself readily for use as a textbook, being somewhat monotonous. Its chief function is to incite the administrator, the supervisor, and probably the classroom teacher to check up on their respective performances in terms of efficiency.

We probably shall not be in agreement with the authors at all points. This does not detract from the value of the book, since its aim seems to be to raise to a definite plane of consciousness many educational problems now too commonly neglected. For the immature student, rapid reading of the book, followed by discussion under the guidance of a competent teacher would be better than private study. To profit most from this book, the reader should have a rather good overview of the general field of education. This applies particularly in the matter of aims and functions of education on the various levels, of standards, measurement and tests, of curricula, methods of procedure and personnel.

T. H. SCHUTTE
Woman's College of Alabama.

UHL, WILLIS L., AND OTHERS. *The Supervision of Secondary Subjects.* New York City: D. Appleton & Company, 1929. xvi + 673 pp.

This promises to be one of the most helpful and suggestive of recent books in the field of secondary education. It covers the ground of supervision in such a way as to aid not merely the general supervisor and the department head, but also the individual teacher. It is written by specialists who have earned the right to speak with authority.

In chapter one the editor, Dean Willis L. Uhl of the University of Washington,

describes the task of the supervisor as that of "knowing the place of each department in a system of education" and "guiding the solution of many problems which few specialists can either see or solve." In this connection he argues that it is the obligation of the supervisor (1) to formulate a system of general educational objectives, (2) to select and encourage significant teaching in pupil activities, (3) to assist as "expert counselor in the selection and organization of content in all departments," (4) to plan for dependable teaching procedures, (5) to formulate valid standards of achievement, and (6) to devise adequate measures of accomplishments.

With these major lines of supervisory activity laid down, the editor turns the discussion over to the specialists. On the whole these specialists, Charles J. Pieper in Natural Science, John R. Clark in Mathematics, Mason D. Gray in Latin, Robert D. Cole in the Modern Languages, S. A. Leonard in English, H. C. Hill in the Social Studies, E. G. Blackstone in Commercial Education, Bernice Dodge and Ruth Henderson in Home Economics, A. H. Edgerton in Industrial Arts, Ethelwyn Bradish in Art, Edgar B. Gordon in Music, Ethel Perrin in Health Education, and Hannah Logasa in Library Service treat their subjects from the points of view emphasized by the editor. There results a compendium of practice which is at once helpful and challenging. Teachers will read with profit the chapters dealing with the subject of their major interest and the supervisors of secondary education will welcome a concrete approach which contrasts with the barren formalism characteristic of much contemporary discussion of supervisory technique.

While it is true that the editor emphasizes the fact that a supervisor must possess a clear grasp of secondary school objectives and also the importance of the coördinating function of the general supervisor, it can still be said that the book under review deals lightly with a fundamental conception of the purposes of secondary education. The objectives mentioned by Uhl in his first chapter are by no means peculiar to secondary education. Nowhere in the book is a philosophy of secondary education clearly enunciated. If it exists it is tacitly assumed, and consequently the book fails to make explicit a criterion which will operate at once in the organization of curricula and the selection of subject matter and procedures within the special subjects of the secondary school. Significant and valuable as the book truly is, it makes clear that the reorganization movement has failed thus far to bring forth a philosophy of secondary education. And just because of the widespread influence they will exercise, one regrets that the authors have not faced their problem in the large as well as in detail.

V. T. THAYER
Ethical Culture School, New York City

CURTIS, F. D. *A Synthesis and Evaluation of Subject-Matter Topics in General Science.* New York: Ginn and Company, 1929. 82 pp.

The study by Mr. Curtis represents a type which would be extremely helpful in every field at the present time. We need to accumulate curriculum objectives determined during the past ten years of intensive study, first, for the purpose of consolidating the valuable results, and second, to discover the most useful methods of determining objectives in each field. The purpose of Mr. Curtis' study was first, to determine the topics of general science most commonly found in a selected number of textbooks, curriculum investigations and courses of study; and second, to determine the relative import-

ance of these topics. The author selected 3 courses of study, 3 analyses of the contents of textbooks, 7 analyses of the scientific interests of children, 3 analyses of science in newspapers and magazines, and 2 unclassified studies—a total of 18 sources.

For the purpose of tabulating the frequency of occurrence of each topic, the sources were divided into six groups. Every topic was given a rank within its source in accordance with the measure of relative importance assigned to that particular topic in the source. Each topic in sources which did not assign numerical values was given a rank equal to the average of all the ranks included by that group. Each group of sources was then given a relative value. The value of a topic, then, in each group was the product of the value of the topic in the source and the relative value of the group of sources. The final aggregate value of any topic, therefore, was the sum of the values of that topic in the six contributing groups of sources. The result of the study consists of particularly detailed lists of topics and sub-topics, together with a measure of relative values expressed in percent.

The chief strength of Mr. Curtis' study lies in the reduction of the content of general science to minute detail. It is an extremely convenient source of curricular and textbook material. The chief limitation is the burdensome statistical calculation which in the end defeated the second of his two purposes. A superficial examination of the final table shows that 1706 out of about 1850 topics rank somewhere between 45 and 65 percent and that very few topics rank very low or very high. In general, therefore, the relative importance of the topics in Mr. Curtis' table has very little significance. This can be accounted for by the arbitrary manner in which he treated his data. It is erroneous to assign a small value to a topic when there is no evidence in the study for it. It is possible that it was the force of the data of the original study to assign a topic a zero value or conversely, if the study had been complete, it may have had a maximum value. The accumulation of these unearned values is responsible for the lack of spread in the final table. Besides, the cumbersome data which Mr. Curtis heroically tried to manipulate were too varied in nature and in scope to be conveniently comparable.

HENRY HARAP
Western Reserve University

ELSBREE, WILLARD S., HALSEY, H. R., AND ELSBREE, ELIZABETH S. *The Teacher's Handbook.* New York: Bureau of Publications, Teachers College, Columbia University, 1929. viii + 289 pp.

Since the field of education extends as far out as "out," goes as deep down as "down," and reaches as far up as "up," a handbook comprehensive enough to cover the subject must necessarily be of such volume as to assume the proportions of a well-stocked library. The authors of this little book containing 1012 questions and their answers on the various phases of education have fairly stated that "no claim is made by the authors that the knowledge called for constitutes the *sine qua non* of education." It is further stated that "each question was carefully evaluated and was included only if the answer was felt to be worth knowing." Since the technique of evaluating is not given, are we to infer that it was based entirely on feeling?

Reading through the book, one is struck with the apparent pertinence of the questions to the generally taught phases of education; and one should expect this,

no doubt, since the questions were chosen largely from outstanding books on education, and from state examinations for the certification of teachers. In general it appears that the book is much more valuable as a catechism in education than a reference book, although this objection, if it may be termed such, is partly overcome by the fine list of 278 references, referred to by number at the end of the answers, and given completely at the end of the book. Since the nature of the book is catechismic, it appears to the writer that its information would be much more readily absorbed if the answers were given immediately after each question.

The book is well bound, makes a pleasing appearance, and should be of real value to teachers wishing to get an extensive overview of the whole field of education.

W. K. WILSON

Ohio State University

GREGORY, C. A., AND RENFROW, O. W. *Statistical Method in Education and Psychology.* Cincinnati: C. A. Gregory Company, 1929. 228 pp.

"The authors of this text," according to their statement in the preface, "have striven, particularly, to make clear the nature and computation of statistical measures which appear in introductory textbooks on the subject." In fact, they have probably commented upon a larger number of different statistical indices than have the authors of any other introductory textbook of equal length. In addition to the usual measures of average, the harmonic mean and the geometric mean are each given several pages of discussion; and, in the chapter on measures of correlation, more space is given to a group of advanced statistical concepts, including regression lines, correlation ratios, corrections for attenuation, mean square contingency, coefficients of association and of colligation, and bi-serial correlations, than to the product-moment coefficient itself. Although the book is obtensibly intended for the use of "non-mathematical students in education and psychology," the authors seem to have been constantly afraid that some expert statistician might criticize them for having failed to mention his pet statistical concept.

It is unfortunate that the authors did not limit their offerings to those matters which could be fully presented and actually mastered in a one-semester course, or else lengthen the book sufficiently to give adequate discussions, illustrations, and exercises for each operation recommended. After declaring their intention of emphasizing the inductive method of approach, they found themselves so crowded for space that they could in most cases do little more than state the meaning of a term and give an illustration of its use. The application of a new principle or technique to a variety of situations in order to make more probable its transfer to other situations has been neglected repeatedly. Definitions and discussions have in some instances been over-simplified and thereby made inaccurate. The book exhibits frequent evidences of a lack of editorial supervision, and the binding of the reviewer's copy exhibits marked signs of disintegration after a single gentle perusal.

One of the best features of the book is an introductory three page review of certain mathematical principles. The authors are also to be commended for introducing several unusual graphic interpretations of statistical measures. An adequate table of contents, a glossary of statistical terms, an index, and the customary table of squares and square roots add much to the practical value of the volume.

M. R. TRABUE

University of North Carolina

RESEARCH ABSTRACTS and BIBLIOGRAPHIES

Address all communications relative to research abstracts and bibliographies to Editor A. S. Barr, University of Wisconsin, Madison, Wisconsin.

LAZAR, MAY. "Diagnostic and Remedial Work in Arithmetic Fundamentals for Intermediate Grades," *Bureau of Reference, Research and Statistics,* Board of Education of the City of New York, 1928, 204 pp.

A valuable investigation setting forth the results of a diagnostic and remedial program with a normal group of forty-three 6A pupils extending over a period of 5 months. The bulletin provides a discussion of the group experimented upon, the tests used, methods of diagnosis, remedial instruction, case studies, measures of improvement, and a summary of conclusions. There is a list of selected practice materials and a selected bibliography on the teaching of arithmetic in the appendix.

LEAL, MARY A. "Physiological Maturity in Relation to Certain Characteristics of Boys and Girls." Philadelphia: University of Pennsylvania, 1929, 118 pp.

This bulletin contains a study of physiological maturity in relation to certain characteristics of boys and girls, based on data secured in the New Britain, Connecticut, Public Schools, grades two through twelve. The bulletin presents a summary of previous investigations, how boys and girls mature stage by stage, how boys and girls progress through school stage by stage, and the personality traits characteristic of different stages of physiological maturity.

NIFENECKER, EUGENE A. "Diagnostic and Remedial Work in Arithmetic Fundamentals," Educational Measurements for the Class Teacher, *Bureau of Reference, Research and Statistics,* Board of Education of the City of New York, (September, 1929), 15 pp.

A summary statement of methods of diagnosis in arithmetic with a selected bibliography on the teaching of arithmetic. This bulletin is a very valuable one for classroom teachers.

OWENS, ALBERT ALEXANDER. "The Behavior-Problem Boy: A Socio-Educational Survey." Philadelphia: University of Pennsylvania, 1929, 188 pp.

This survey comprises a study of three groups of pupils: first, boys who entered the William A. Lee School in Philadelphia during the year 1926 and 1927; second, boys who were dropped from the register during the same year; third, boys whose names were dropped during the years 1920 to 1922. There were in these groups 383, 377, 257, and 356 respectively, making a total of 1,373 boys about whom data were secured. Besides these data pertaining to factors characteristic of the behavior-problem boy, a review of previous investigation is presented.

Research Bulletin of the National Education Association, "The Principal as a Supervisor," Washington: Research Division of the National Education Association, VII, (November, 1929).

This bulletin contains material upon the nature of supervision, its development, how to proceed in supervision, the importance of knowing the community, the parents, the teachers, and the pupils, methods of classroom visitation, methods of individual conferences, teachers' meetings, demonstration teaching, and pupil diagnosis and adjustment.

"Sixth Annual Conference on Elementary Supervision," *Bulletin of the School of Education, Indiana University*, VI, (September, 1929), 72 pp.

This bulletin contains papers by Melvin E. Haggerty on "Our New Knowledge of Personality," by Clem O. Thompson on "Supervision of Beginning Teachers," by Lenna E. Smock on "The Relative Difficulty of the Reading Content of Intermediate Grade History Textbooks," and other papers.

WRIGHT, WENDELL WILLIAM. "The Development and Use of a Composite Achievement Test," *Bulletin of the School of Education, Indiana University*, V, (January, 1929), 90 pp.

This bulletin presents data pertaining to the validity, reliability, and weighting of the Indiana Composite Achievement Test, with a summary of results as derived from four types of rural schools. The Indiana Composite Achievement Test, when combined with ratings for industry by the method of multiple correlation, correlated .827 ± .01, with success in the ninth grade as measured by average school marks.

EXPERIMENTAL STUDIES REPORTED IN CURRENT ISSUES OF OTHER MAGAZINES

EDUCATIONAL ADMINISTRATION AND SUPERVISION

ANDERSON, EARL W. "Extra-Curricular Duties," *Educational Research Bulletin*, Ohio State University, VIII, (October 9, 1929), 315-317.

Extra-curricular duties often require a considerable part of the teacher's time, energy, and initiative. The 110 graduates from Ohio State University answering the questionnaire used in this investigation reported 53 different types of activities in which these graduates engaged. The chances are more than four to one that beginning high-school teachers will be responsible for extra-curricular duties. The author concludes that definite preparation for the supervision of these functions should be required as a part of the teacher-training course.

BUTSCH, RUSSELL L. C. "A Comparative Study of the Effects of Different Types of School-Building Ventilation on the Health of Pupils. III," *The Elementary School Journal*, XXX, (November, 1929), 208-217.

The New York State Commission on Ventilation seemed to be of the opinion that greater air flow in mechanically ventilated rooms is in part responsible for the greater amount of respiratory illnesses reported for buildings so ventilated. The present study relates to this problem. The author concludes that the physical conditions of the room in general, including temperature, humidity, rate of air flow, and general cooling effect, have no significant influence upon absences due to respiratory illnesses.

FRASIER, CLARK M. "Intelligence as a Factor in Determining Student Teaching Success," *Educational Administration and Supervision*, XV, (November, 1929), 623-629.

The author concludes that intelligence as measured by standard intelligence tests is not an important factor in determining success in student teaching. Low correlation is attributed to the fact that those with less intelligence than is necessary for teaching success have already been weeded out in high school and that success in student teaching depends upon other factors than intelligence.

KEFAUVER, GRAYSON N. "The Validity of Bases for Forming Ability Groups," *Teachers College Record*, XXXI, (November, 1929), 99-114.

After summarizing briefly the investigations of Sorenson, Pickell, Brooks, Fretwell, Kelley, and Dvorak, the author presents the results of his own investigation, which is based upon a study of ability grouping in the seventh grade of the junior high schools in Fresno, California. The author concludes from the evidence obtained in his study that, whatever composite measure is used for ability grouping, it should include: (1) a composite of marks obtained in the earlier grades or a rating of capacity by the teachers in these grades, or both; (2) the intelligence quotient or the score on a mental test; and, when it is expedient to extend the program of testing, (3) a general achievement test or a composite of a number of special achievement tests.

RUSSEL, F. O., AND TWINING, PAUL E. "A Comparative Study of the Academic Efficiency of Certain Urban and One-Teacher Rural Schools," *University of Kansas Bulletin of Education*, II, (October, 1929), 18-20.

The achievement of fourth-grade and eighth-grade rural school pupils in 14 scattered districts in Bourbon County were compared with similar groups of urban students in the Fort Scott public schools, Kansas. When the groups were equated for intelligence, chronological age, and grade in school, the results in one grade (the fourth) favor urban students. The results in the eighth grade, however, favor rural students. The differences were not great in either case.

SORENSON, HERBERT. "Why Teaching Success Does Not Correlate Highly with Measured Intelligence," *Educational Administration and Supervision*, XV, (November, 1929), 602-606.

The author attributes the low correlations ordinarily found between intelligence and teaching success to errors of measurement and the selective character of the groups measured.

PSYCHOLOGY OF LEARNING

HANSEN-TYBJERG, CHR. "Differences in the Talents and Abilities of Young People," *The Journal of Applied Psychology*, XIII, (October, 1929), 451-468.

A study made in Copenhagen, Denmark, of 238 printers' apprentices who were given a number of tests, including dictation, manuscript reading, proof reading, memory, advertisement and title setting, and arithmetic. The best pupil proved to be twice as quick as the slowest in the first series of tests, and two and one-third times quicker than the worst in the other essays. Similar differences were noted in spelling, reading, proof reading, etc.

HULL, CLARK L. "A Functional Interpretation of the Conditioned Reflex," *Psychological Review*, XXXVI, (November, 1929), 498-511.

A readable article explaining learning in terms of the conditioned reflex. The author recognizes a primary or excitatory period and a secondary or inhibitory period in all learning. There are at least four discernible aspects of the primary phase of the conditioned reflex, namely, re-integration, patterned stimuli, irradiation, and anticipatory tendencies. Corresponding to the four aspects of the excitatory phase, each to each, there are four parallel inhibitory phases, namely, experimental extinction, conditioned inhibition, differential inhibition, and the inhibition of delay. Learning is explained in terms of these psychological processes.

SHAW, LENA, AND CRUMPTON, CLAUDIA E. "The Attitude of the Child in Matters of Skill," *The Elementary School Journal*, XXX, (November, 1929), 218-222.

The study here reported was preceded by a series of studies covering several years of investigation of the handwriting of boys and girls in Detroit. Handwriting is not a part of the Detroit junior high-school curriculum. Some pupils write well, others poorly. Instead of requiring those who write poorly to do special work, the matter is handled through a handwriting club in the English department. The author concludes that attitude is a potent factor in the development of skill.

THURSTONE, L. L., AND ACKERSON, LUTON. "The Mental Growth Curve for the Binet Tests," *The Journal of Educational Psychology*, XX, (November, 1929), 569-583.

The findings herein reported contradict the hypothetical mental growth curve which is usually shown to have a negative acceleration from birth without an inflection point. From the investigation the authors conclude that the mental growth curve has a positive acceleration up to the general age level of about ten years, with an inflection point somewhere in the age range nine to twelve. The inflection point comes earlier for bright children than for dull children. The absolute variability of test intelligence increases with age until adult intelligence is attained.

EXPERIMENTAL STUDIES OF TEACHING

ALTHAUS, CARL B., AND GILLILAND, JAMES F. "The Value of Instruction in How to Study," *University of Kansas Bulletin of Education*, II, (October, 1929), 3-4.

A master's dissertation dealing with the results of an experiment, relative to the value of instruction on how to study, in four sections of sophomore English in the Hutchinson, Kansas, High School. The experimental group was given definite instruction once a week for 26 weeks in how to study, based upon a discussion of the rules presented in Whipple's "*How to Study Effectively*," Sandwick's "*How to Study and What to Study*," and Kornhauser's "*How to Study*." The conclusion reached in this investigation is that instruction in how to study does not improve school work.

STABLER, DEWEY A. "The Relation between the Civic Information Possessed by Ninth-Grade Pupils and their Practices in Citizenship," *The School Review*, XXXVII, (November, 1929), 697-706.

A study of the relationship between the civic information possessed by 120 ninth-grade pupils and their practices in citizenship. A coefficient of correlation of .41 was found between the total average scores on the tests of civic information and the number of civic deficiencies

noted in the pupils studied. The author concludes that the possession of a greater amount of civic knowledge aids pupils in their school citizenship practices.

STONE, CHARLES A. "Effective Methods of Teaching Subtraction of Signed Numbers," *School Science and Mathematics*, XXIX, (November, 1929), 850-858.

A very interesting article on methods of teaching the subtraction of signed numbers. The methods discussed are (a) the method involving the use of the number scale, (b) a tabular method recommended by Wells and Hart, (c) the Austrian method, (d) a method presented by Nyberg, (e) a method presented by Schultze, and (f) a method presented by McCaughan. The article appears to be one of considerable practical value to the classroom teacher.

WILLIAMS, RALPH R. "Extensive Reading Versus Intensive Study of Literature," *The School Review*, XXXVII, (November, 1929), 666-678.

A study of extensive reading versus intensive reading of literature, involving some 200 ninth-grade classes in English. The investigation is of the experimental sort. The author concludes that there is no significant difference in the comprehension of literature attained by the intensive-study classes and extensive-study classes. Neither was there any significant difference in the ability of the two groups to remember the details of the selections read. Considering the time spent, the extensive reading method is more effective in achieving the aims of instruction in literature than intensive-study methods.

CURRICULUM RESEARCH

MINGO, JANE. "A List of Materials and Books Suitable for Use in Remedial Reading," *Teachers College Record*, XXXI, (November, 1929), 137-147.

A list of materials and books suitable for use in remedial reading.

TURNEY, AUSTIN H., AND UMMEL, MAUD. "An Analysis of Frequency of Error in Grammar and Sentence Structure Among Selected Junior High-School Pupils," *University of Kansas Bulletin of Education*, II, (October, 1929), 14-17.

A study to determine the frequency and kinds of errors made by junior high-school students in free writing. The results of the investigation are presented in tabular form.

EDUCATIONAL TESTS AND MEASUREMENTS

ANDERSON, ERNEST M. "A Study of the Reading Ability of College Students," *University of Kansas Bulletin of Education*, II, (October, 1929), 5-9.

An abstract of an investigation by E. M. Anderson on individual differences in the reading ability of college students, based upon a bulletin published by the University of Missouri, 1928. The evidence indicates that there are extreme individual differences in the reading ability of college students, that there is a high correlation between intelligence and reading ability, and that many students cannot read sufficiently well to do the assignments made in their course work.

CURETON, EDWARD E. "Computation of Correlation Coefficients," *The Journal of Educational Psychology*, XX, (November, 1929), 588-601.

The author presents a method of correlation involving the sum of the squares of the sums of gross scores and certain checking formulae especially adapted to the use of machine calculation.

RESEARCH NEWS and COMMUNICATIONS

Address all research news and communications to Editor V. T. Thayer, Ethical Culture Schools, 33 Central Park West, New York City.

FOUNDATIONS FOR RESEARCH

A survey recently completed by the Office of Education indicates that there are fifteen separate and independent boards and foundations dispensing funds for the promotion of education in the United States in addition to Federal and State school boards. Some of the interesting features of this survey as summarized by the *New York Times* for September 16, are as follows:

"The General Education Board, from its foundation in 1902 to June 20, 1928, appropriated $17,087,062.74. Of this sum $12,462,000 represents appropriations from the principal and $5,025,062.74 from income.

"The Rockefeller Foundation for 1927 disbursed from income and capital for medical education alone $4,097,343. The Laura Spelman Rockefeller Memorial during 1928 appropriated $38,082,058.34 for educational, scientific and charitable purposes and contingencies. On Jan. 3, 1929, the Laura Spelman Rockefeller Memorial was consolidated with the Rockefeller Foundation.

"The Carnegie Corporation of New York authorized appropriations of $2,041,250 for the fiscal year for library service, adult education, the arts, educational studies, research, publications and general interests. The Carnegie Foundation for the Advancement of Teaching in the year ended June 30, 1927, appropriated $2,685,025.26 for retiring allowances and pensions in institutions and for individuals, for pension studies, administration and publications, and $1,501,349.35 for studies of legal and dental education, graduate instruction, other studies and college athletics.

"The John F. Slater Fund, through its educational committee, for the year 1927-28, appropriated $69,050 for assistance in paying the salary of a professor in the English or science department in each of twenty colleges.

"The Jeanes Fund, for the improvement of negro rural schools, coöperated during the season ended June 30, 1927, with public school boards and superintendents in 306 counties in fourteen states, appropriating $106,547.

"The Phelps Stokes Fund, incorporated by the State of New York in 1911 with a capital of $1,200,000; 'for educational purposes in the education of negroes, both in Africa and in the United States; North American Indians and needy and deserving students;' has founded fellowships in the Universities of Georgia and Virginia for the study of the negro problem.

"It has established a special fund at the George Peabody College for Teachers, at Nashville, Tennessee, to enable students and teachers there to visit colored schools and observe progress.

"The American Field Service Fellowships for French Universities, which is administered by the Institute of International Education, awarded nine fellowships for the year 1928-29.

"The Commission for Relief in Belgium, the Education Foundation (Inc.)

and the Foundation Universitaire during 1927 contributed $150,000 for the completion of the Louvain Library. The foundation authorized a contribution of 9,500,000 Belgian francs (about $1,319,500) to complete the main group of the buildings of the University of Brussels.

"The Kahn Foundation for the Foreign Travel of American Teachers organized in 1911, supports a single fellowship of $5,000 yearly.

"The Commonwealth Fund during the fiscal year ending September 30, 1928, appropriated $2,083,621.80 for public health and child welfare.

"The Julius Rosenwald Fund, incorporated, in 1917 under the laws of Illinois, appropriated for the fiscal year ending June 30, 1928, for charitable, scientific, educational and religious purposes $364,831.21.

"The Payne Fund, organized in September, 1927, provides funds for the study of juvenile readings and appropriates the income on about $2,000,000."

How Effective is Specific Training in Preventing Loss Due to the Summer Vacation?

Kirby[1] has shown the effect of specific training in addition and division. He found that it required of addition, approximately, thirty minutes of drill, in the spring, to bring the group to the same level of efficiency as it had attained in the fall by seventy-five minutes of drill. In division, "more than three-fifths as much practice was required to regain the standing reached in the experiment as was required at that time to reach it." Nelson[2] in a recent study has shown that it required considerable time to attain the level of efficiency as reached in the previous spring in most subjects. Bruene[3], has likewise shown the detrimental effect of the summer's vacation upon the fundamentals of arithmetic, reasoning problems and spelling, while there is an increase in efficiency in reading. This was true in grade four. In grades five and six there was an increase in efficiency in reading and nature study, while there was a loss of efficiency in the fundamental processes, problem solving, history, language usage and spelling. Other writers, such as Elder[4], Kramer[5], and Patterson[6], have also shown the effect of the summer's vacation upon efficiency in school subjects.

The present study is concerned with two questions: (1) The effectiveness of specific training in preventing loss of efficiency due to summer vacation; (2) The effect of the summer vacation upon efficiency of the two groups.

This study was carried on in a city in Southeastern Kansas. Two sixth-grade classes were used. The same teacher taught both groups in the subjects considered in this study. The following tests were given to both groups on May 11, May 25, and September 4: Compass Survey Test in Arithmetic, Form A; Thorndike-McCall

[1] Kirby, Thomas J. *Practice in the Case of School Children.* Teachers College, Columbia University Contributions to Education, No. 58, 1913.

[2] Nelson, M. J. "How Much Time Is Required in the Fall for Pupils in the Elementary School to Reach Again the Spring Level of Achievement?" *Journal of Educational Research,* xviii, (Nov. 1928), pp. 305-308.

[3] Bruene, Elizabeth. "Effect of Summer Vacation on the Achievement of Pupils in the Fourth, Fifth and Sixth Grades." *Journal of Educational Research,* xviii, (Nov. 1928), pp. 309-314.

[4] Elder, H. E. "The Effect of Summer Vacation of Silent-Reading Ability in Intermediate Grades." *Elementary School Journal,* xxviii, (March, 1927), p. 541.

[5] Kramer, G. A. "Do Children Forget During Vacation?" *Baltimore Bulletin,* vi, (December, 1927), pp. 56-60.

[6] Patterson, M. V. W. "The Effect of Summer Vacation on Children's Mental Ability and on their Retention of Arithmetic and Reading." *Education,* xxxvi, (Dec., 1925), pp. 222-228.

Reading Scale, Form 8; and, Otis' Reasoning Test in Arithmetic, Form A. To one group, which we shall designate as Y, special training was given in the fundamentals of arithmetic, silent reading, and problem solving. This training lasted for two weeks. The training in the fundamentals of arithmetic consisted in administering four diagnostic tests, one in addition, subtraction, multiplication, and division. Then the weaknesses revealed in these diagnostic tests were followed by remedial teaching, which consisted in using the Economy Remedial Exercise Cards. Ten minutes practice daily for the period was devoted to this work.

In reading, mimeographed material similar to that in the scale was used, daily; specific questions being asked on the material read. Also, three forms of the same test were used. Part of each recitation period consisted in revealing to the pupils the correct answers to the questions asked, and they were led to understand that the answers were to be found in the material read.

In problem solving, an analysis was made of the last ten problems in the test, and then five similar problems were made for each of these ten problems given in the test. The pupils were asked to pay attention to the following things: (1) to understand each word in the problem; (2) to determine what is given in the problem; (3) to determine what is required; (4) to select the different processes to be used in solving the problem; and, (5) to lay out the solution by asking: (a) what is given? (b) what am I to find? (c) how am I to use what is given in order to find out what I should prove? (d) how can I check my answer to prove that it is correct?; and, (6) how to check the answer. For each problem the pupils were required to have their plans checked as given in (5) above before they were permitted to proceed in solving the problem. After the plans were approved, they solved the problem, and then were required to check their answers.

After this two weeks' training, the same form of the same tests were given to both groups. This was during the last week of school before the summer vacation began. None of the pupils in either group attended summer school. On September 4, the same form of the same tests were given again. Neither the teacher nor the pupils knew that the tests were to be administered when school convened after the summer vacation. A summary of the results is given in the tables following:

It will be seen that group Y was superior to group X in the initial test in the fundamentals of arithmetic by 1.09 problems, and in problem solving ability .23 problems, while group X was superior to group Y in reading ability by .15 questions.

By two weeks of training in the fundamentals of arithmetic, group Y increased in efficiency 8.10 problems while group X increased only .21 problems. In reading, group Y increased 4.25 questions, while group X increased only .29 questions. In problem solving group Y increased 2.43 problems, while group X increased only .39 problems. Table II, will show gains made in A.M.'s in the various subjects during the two weeks of training, and the index of significance of these gains.

Table II shows that group Y increased its gain over group X in the fundamentals of arithmetic by 7.86 problems, in reading by 3.96 questions, and in problem solving by 2.04 problems.

We now come to the second question,

TABLE I

Group—	N	Fundamentals of Arithmetic			Reading			Problem Solving		
		May 11	May 25	Sept. 4	May 11	May 25	Sept. 4	May 11	May 25	Sept. 4
Y.	40	27.88	35.98	33.42	18.50	22.75	23.03	8.70	11.13	10.90
X.	38	25.79	26.00	22.31	18.65	18.94	19.66	8.47	8.86	8.71
Difference.		1.09	9.98	11.11	.15	3.81	3.37	.23	2.27	2.19

TABLE II

Fundamentals of Arithmetic

Group	A.M. Gain	P E A.M.	S.D.	P E diff	Index of Significance	Difference Due to Specific Training
Y.	8.10	.973	9.13	1.252	6.302	
X.	.21	.782	8.04			7.86

TABLE III

Reading

Group	A.M. Gain	P E A.M.	S.D.	P E diff	Index of Significance	Difference Due to Specific Training
Y.	4.25	.794	7.45	1.123	3.526	
X.	.29	.795	7.28			3.96

TABLE IV

Problem Solving

Group	A.M. Gain	P E A.M.	S.D.	P E diff	Index of Significance	Difference Due to Specific Training
Y.	2.43	.401	3.76	.526	3.878	
X.	.39	.341	3.12			2.04

TABLE V

Loss of Gain by Vacation	Fundamentals of Arithmetic	Reading	Problem Solving
	Gain—Loss	Gain—Loss	Gain—Loss
Group X	3.69	.72	.05
Group Y	2.56	.28	.23
Difference.	1.13	.44	.28

TABLE VI

Group	Fundamentals of Arithmetic			Reading			Problem Solving		
	Gain	Loss	Same	Gain	Loss	Same	Gain	Loss	Same
Y.	2	35	3	16	12	12	9	16	15
X.	1	36	1	25	8	5	11	20	7

i. e., the effect of the summer vacation upon the two groups. Were the two groups equally affected? Table V reveals that the fundamentals of arithmetic group X lost 3.69 problems while group Y lost only 2.56 problems, or a difference of 1.13 in favor of the Y group. In reading, however, group X gained .72 questions while the Y group gained only .28 questions or a difference of .44 in favor of the X group. In problem solving group X gained .05 problems, while group Y lost .23 problems.

For the Y group, 35 pupils lost in efficiency during the vacation, 3 remained the same, while 2 gained in efficiency in the fundamentals of arithmetic. In reading, 16 pupils gained in efficiency, 12 remained the same, while

12 lost in efficiency. In problem solving, 9 pupils gained in efficiency, 15 remained the same, while 16 lost in efficiency.

For the X group, 36 pupils lost in efficiency, one gained, and one remained the same in the fundamentals of arithmetic. In reading, 25 increased in efficiency, 5 remained the same, while 8 lost. In problem solving, 20 pupils lost in efficiency, 7 remained the same, while 11 gained. The results are given in Table VI.

The ten poorest students in problem solving in group X lost in efficiency during the vacation, while in the Y group the poorest students gained and the three next remained the same.

It might be asked, just how much value was the two weeks' specific training in the two subjects? The results are given in Table VII. The results in the fundamentals of arithmetic are only approximate, because the test norms do not express scores by months, so it was necessary to approximate the gains in this instance.

The loss or gain due to summer vacation may also be expressed in months, as given in Table VIII.

The following implications may be drawn from this study: first, that two weeks' specific training is productive of greater efficiency in the three subjects considered, the greatest gain being made in reading, which was equivalent to about twenty months. In problem solving, the increase was seventeen months, while in the fundamentals of arithmetic, it was seven months. Both groups were rather low in efficiency in the fundamentals of arithmetic at the beginning of the experiment, and even by specific training were not able to attain the norm established for the six high grades. In reality these groups were practically seventh graders. The writer would not have it implied, however, that such training would be productive of equal increments in efficiency if carried on indefinitely. Second, that the loss for group Y exceeded that of group X in problem solving. The greater loss may have been due to the shortness of the training period. Third, that in the fundamentals of arithmetic, where mere skill is involved, group Y did not lose to the same degree as group X. Fourth, that in reading, both groups increased in efficiency over the vacation, but group X had the greater increase. Group Y may have attained its approximate maximum of efficiency during the training period. Fifth, that the specific training acted as a sort of buffer to prevent loss of material previously learned; for in an analysis of the errors made before and after the vaca-

TABLE VII

	Fundamentals of Arithmetic	Reading	Problem Solving
Group	Gain	Gain	Gain
Y.	7 months	20 months	17 months
X.	2 months	1 month	1 or less
Difference due to training	5 months	19 months	16 months

TABLE VIII

	Fundamentals of Arithmetic	Reading	Problem Solving
Group	Gain—Loss	Gain—Loss	Gain—Loss
X.	3.5	3	0
Y.	2.5	2	1.5
Difference	1.0	1	1.5

tion period it was found that 77 percent of the errors were identical.

The writer is planning to repeat the experiment in grades two, four, and six, and to use two hundred pupils in each group for each grade. Possibly a larger sampling may produce different results.

Pittsburg, Kansas L. D. MORGAN

THE VOCABULARY OF PLANE GEOMETRY

Five years ago studies were made of the "Vocabulary Burden" of certain text-books in different subjects.[1]

Recently[2] a number of texts in plane geometry and in elementary algebra were studied by practically the same methods. From a thousand words (taken at regular intervals throughout the whole book) calculations were made of: (1) the "vocabulary range" (the number of different words); (2) the number of words of "zero index" (the words not found in E. L. Thorndike's *Teacher's Word Book*); and, (3) the median index (using Thorndike's "credit-number" for each of the words found in his book). The entire calculation was also made for another sampling of one thousand words from each textbook.

The purpose of the present paper is three-fold:—(1) to make a few comments on the interesting data for the plane geometry texts studied by Remmers and Grant; (2) to apply their method to Forms A and B of a published new-type test in plane geometry that is widely used; and (3) to apply a different method to an intensive study of the words of zero index in the plane geometry vocabulary.

Anyone who is familiar with the recent and the present tendencies in the teaching of plane geometry will note that it is the earliest (1907) of the six texts on this subject that has the lowest median index ("about 25"), the largest vocabulary range (301-312), and the largest number of words of zero index (54-57); while the 1925 text has the highest median index ("about 60") and the 1923 text has the smallest vocabulary range (220-225) and the smallest number of words of zero index (23-25). This simply confirms our opinion that there has been a marked tendency in the last twenty years to simplify the vocabulary of our subject.

The writer's study, of the new-type test, gave for Forms A and B respectively: Vocabulary range, 184 and 192; number of words of zero index, 27 and 28; and, median index, 151 and 149. These compare favorably with the most recent texts examined by Remmers and Grant. But the writer raises the question whether it is fair to compare texts and standardized tests. Is it not necessary to list *every* word of zero index in a test which is to be taken by candidates coming from schools that have used various textbooks? If a new word is found in a textbook there is an opportunity for explanation; but if a candidate meets unfamiliar words in a test he is placed at a disadvantage. In such a case it is quite possible for a pupil to receive a rating which in no way represents his knowledge of the subject or his power to apply that knowledge.

As the 10,000 words in the *Teacher's Word Book* are those that "were found to occur most widely in a count of over four and a half million words," it is safe to assume that words not found there may be considered as belonging to a technical vocabulary. What are the indispensable words of the technical vocabulary of plane geometry? This ques-

[1] Lively, Bertha A., and Pressey, S. L. *Educational Administration and Supervision*. IX, (October, 1923), pp. 389-398.

[2] Remmers, H. H., and Grant, A. "The Vocabulary Load of Certain Secondary School Mathematics Textbooks." *Journal of Educational Research*. October, 1928, pp. 203-210.

tion was not so important twenty years ago, when the classics played an important part in the secondary-school curriculum, as it is to-day when many of our high-school pupils can profit by a course in geometry though they have had little or no linguistic work except in English.

In the search for a minimum technical vocabulary for plane geometry, it seemed advisable to make lists of the words of zero index in the lists of theorems and constructions of: (1) the College Entrance Examination Board syllabus (major requirement, Document 198); (2) the report of the committee on the "Reorganization of Mathematics in Secondary Education;" and, (3) the syllabus (in plane geometry) issued by The University of the State of New York Press, 1927. The first source yielded the following list of twenty-five words:

are	perimeter
apothem	polygon
bisect	quadrilateral
bisector	ratio
circumscribe	secant
congruent	segment
convex	semicircle
equidistant	supplementary
hypotenuse	tangent
intersect	transversal
locus	trapezoid
median	vertex
parallelogram	

The second source gave only one additional word, *decomposed*. As this occurred only once and was used instead of the more usual word *separated* (or *divided*), it may safely be disregarded. The third source gave no new words.

As the first two sources are the published reports of separate committees of recognized organizations of national reputation, and as the third source is a published report of the educational system of the State of New York, this list of twenty-five words has deep significance. It is evident, of course, that where explanations could be given, this list does not contain every word that might be used in proofs or definitions in a textbook.

The really remarkable feature of this list is the absence of certain words, such as *concurrent, dodecagon, equilateral, equiangular, mutually equiangular, isosceles, rhombus, subtend,* etc., which were considered indispensable by the older teachers and writers of textbooks, and which are still dear to the hearts of the makers of some of the widely used new-type tests.

Let us compare the vocabulary of Forms A and B of the test already studied with our new list of twenty-five words. In Form A, twenty-three of the hundred questions contain words of zero index not found in our list. *Dodecagon, mutually equiangular, sector,* and *subtend* each appear once. Each of the words *equiangular, isosceles,* and *rhombus* is found in three questions, and *equilateral* is used in ten questions. In Form B, *concurrent, isosceles, octagon,* and *rhombus* each occur once. *Projection* is found in two questions, *sector* and *subtend* each appears in four questions, and *equilateral* appears in five. In all, there are nineteen questions each containing at least one of these words.

What is the effect of this upon progress in the teaching of plane geometry? Take, for example, the inclusion of the word *isosceles*. The old textbooks stated one theorem "In an isosceles triangle, the angles opposite the equal sides are equal." In each of the three lists of theorems discussed above, this theorem is written in the simpler form, "if two sides of a triangle are equal, the angles opposite these sides are equal." If a teacher uses this simpler form, he handi-

caps his pupils if they are required by educational authorities to take a standardized test that contains the word *isosceles*. Even in the case of accepted words there is the possibility of difficulty. For example, more than one test uses the word *circle*, as in the older texts, to mean a "portion of a plane bounded by....", referring to the curve itself as the circumference of the circle. This will confuse the candidate who has been accustomed to the more modern usage, where the curve is called the circle and the word *circumference* is reserved for its numerical measure.

It is plainly evident that the makers of standardized tests that are widely used have it within their power to retard or to accelerate the movement to lighten the vocabulary load in plane geometry.

ELIZABETH B. COWLEY

STUDENTS' ESTIMATES OF THEIR ABILITY AND ACHIEVEMENT

The purpose of the following experiment was to compare (1) students' estimates of their mental ability with the ranking revealed by the Otis Self-Administering Tests of Mental Ability, and (2) students' estimates of their scholastic achievement with the semester marks awarded by the teacher.

Procedure.—Immediately before the final semester examinations 111 sophomore women in general psychology were asked to estimate both their percentile ranking in mental ability and the mark they merited in the course. They were then given the Otis Self-Administering Tests of Mental Ability (Higher Examination, Form A). This experiment was made at the close of the instructor's first semester in the institution. Test papers had not been returned during the course nor had the students at any time after the middle of the semester been made aware of their standing. The students were assured that their estimates would not affect their marks. All of them willingly engaged in the experiment.

Table I clearly shows that it was the less capable students who overestimated their ability most. The eleven whose percentile ranking is not over 30 overestimated their ability by upwards of 36

TABLE I

A COMPARISON OF THE PERCENTILE RANKING IN MENTAL ABILITY AS REVEALED BY THE OTIS SELF-ADMINISTERING EXAMINATION WITH THE RANKING AS ESTIMATED BY THE STUDENTS THEMSELVES

Percentile Rank as per Otis Tests	Estimated P. R. deviates from actual: Agreement	Positive: 1-5	6-10	11-15	16-20	21-25	26-30	31-35	36+	Negative: 1-5	6-10	11-15	16-20	21-25	26-30	31-35	36+
1-10									4								
11-20									3								
21-30									4								
31-40			1		1		1		7								
41-50	1	1	2			1		1	2								
51-60		1	1	1	4	1	2	1	1	1	1						
61-70		1	5	2	3	2				1		2					
71-80	3	2	4	4						1	2	1	1				
81-90		1								5	4	3	1	1	1		
91-100										1	5	3	5	1	2	1	

Read Table I: Four persons ranked between 1 and 10 on the Otis tests over-estimated their ranking by more than 36 points; one person ranked between 41 and 50 placed herself correctly, one over-rated herself between 1 and 5 points, two over-rated themselves 6 to 10 points, etc. Similarly the negative side of the table shows the amount of underestimating done.

points and seven of those whose actual ranking is between 31 and 40 did likewise. Those who actually rank from 41 upward show much less tendency to judge themselves too highly. Those who rank above 50 show an increasing tendency to under-estimate their mental status.

TABLE II

THE PERCENTILE RANKING OF STUDENTS AS REVEALED BY THE OTIS TESTS AND THE ESTIMATES OF THEIR CLASS MARKS SHOWN FOR THE VARIOUS LEVELS OF MENTAL ABILITY

Percentile Rank in Mental Ability as per Otis Tests	Estimate of class or term mark by students compared with teacher's awards						
	Same as Teacher	Over estimated			Under estimated		
		+1	+2	+3	−1	−2	−3
1-10			1	3			
11-20		1	2				
21-30		2	1	1			
31-40	1	5	4				
41-50	3	3	1	1			
51-60	6	6	2	1			
61-70	6	7	1		2		
71-80	6	8	1		2		
81-90	7	5	1		4		
91-100	9	7			3		

Table II should be read as follows: Of the students whose percentile ranking in mental ability is between 1 and 10, no one estimated her class mark the same as that awarded by the teacher. The nearest agreement was in the case of one who put her grade two points above what the teacher gave her. Three ranked themselves three points higher than the instructor placed them, while no one put her mark too low. Of those who ranked below 31 in mentality not one judged her work in accordance with the teacher's grade or below it. All but one of the twenty persons whose percentile ranking in mental ability is below 41 thought their class marks would be higher than they were. None with ranking below 61 in mentality expected a mark below that received.

This table clearly reveals that those persons in the lower realms of intelligence tend to overestimate their class merits, while those who underestimate them are found in the higher mental levels. Not only do those of mediocre mentality overestimate most frequently, but the amount of errors is correspondingly greater. The higher the mental ability of the student, the more likely is her judgment to harmonize with that of the teacher, with regard to her worth in the course. When her estimate does vary from the teacher's award, the difference tends to be small.

General Results.—Sixty-four of the 111 students estimated their mental ability higher than the Otis tests revealed it; 43 estimated it lower, and 4 agreed with the Otis ranking. Sixty-two thought their semester marks would be higher than those awarded; 11 expected them to be lower, while in 38 cases the judgment of the student coincided with that of the teacher.

Of those who underestimated their standing in the course, no one varied from the instructor's mark more than one point. (Marks were counted as follows: A = 5; B = 4; C = 3; D = 2; E = 1; F = 0). There was greater error on the part of those who overestimated. Forty-nine of these misjudged by one point, 14 by two points and 6 by 3 points.

Some outstanding facts are revealed by this study. (1) Many students consider their mental ability to be higher than the Otis tests show it to be. (2) Many students place a higher value on their accomplishment than did their teacher. (3) Those who overrate themselves in mentality tend to think too highly of their attainment also. (4) Those who put too low an estimate on their mental ability tend to harmonize more nearly with the teacher's judgment

of their achievement than do those who overrate their intellect. (5) Comparatively few students accurately rate their mentality. (6) Those who overestimated their mental status were less nearly accurate than were those who underestimated it. The median error in the former case is approximately 19.9 points and only 10.8 points in the latter, while the mean error in the former group is 26.7 points, and in the latter, only 12.7 points. (7) Far more students estimate their achievement in harmony with the teacher's rating or below it, than those who rank themselves accurately in mental ability or underrate themselves on this point. This may be because they take greater pride in their native ability than they do in their attainments.

Perhaps we have been too exacting in stating whether or not a student overestimated, underestimated or accurately rated her percentile ranking in mental ability. Since the percentile scale ranges from 0 to 100, an error of five or ten points is probably not serious. Estimating the semester marks, in which the scale runs from "A" to "F," that is from 5 to 0, should not be so difficult. This probably explains why so many more were in agreement with the teacher's semester awards than with the Otis test ranking.

Let us be a little more liberal and regard all who made an error of only five points in estimating their percentile rank of mental ability as being in practical agreement with the findings of the Otis tests. We then have fifteen more students, or a total of nineteen, in agreement with the test rankings. Should we ignore errors of ten points in making percentile ranks we would find a total of forty-four students, instead of four, in practical agreement with the Otis results. There would then be forty-six cases of too high estimate and twenty-one too low.

One may readily conjecture that simple lack of judgment caused persons of inferior mentality to overestimate themselves by large margins. It is not so easy to explain why those of higher endowments tend so largely to underrate themselves. We refer particularly to the fourteen who rank above 71 on the Otis test but underestimate themselves by more than fifteen points. Diffidence may enter in at times, but the writer doubts that it was a factor of any considerable importance in this study. It seems more likely that the work in school (high school and college) has not taxed these brighter students to capacity. They have, therefore, developed poor work habits resulting in mediocre class marks which naturally led them to believe themselves to be persons of ordinary ability.

T. H. SCHUTTE
Womans College of Alabama

A suggestion for your end-of-semester testing

The purpose of a standardized test is to furnish information upon which to direct instruction. This information must be definite for the members of the class and must point out immediate requirements for the development of each pupil through the educational activities of the next semester.

A widely and fully standardized series of tests such as the *Public School Achievement Tests by Jacob S. Orleans* enables you to determine for each pupil the relation of his present achievement to his ability in comparison with American children. Instruction can then be directed to 100% accomplishment for that ability.

We furnish standards for these tests for each of the four 25-percentile groups. These are nation-wide standards determined from over 200,000 pupils quite evenly distributed over the whole country. Their value for instruction is illustrated thus:

> **A pupil in the *upper fifth grade* whose score is 29 or lower on the Public School Achievement Test in Reading is in the slowest 25% of American children in that grade. His instruction should be of the type given to the middle 25% of the pupils in the *upper fourth grade*.**

The table of nation-wide medians and percentiles furnished with the *Public School Achievement Tests* gives these facts for each of ten elementary school subjects and for each half grade from 3B to 8A. Send for the complete report. It is sent *free* upon request.

The thousands of schools which have used the *Public School Achievement Tests* appreciate that these tests are thorough, yet simple in organization; most practical, yet scientifically correct; extensive in the number of subjects covered, yet economically arranged. The subjects covered by the tests are: Reading, Arithmetic Computation, Arithmetic Reasoning, Language Usage, Spelling, Grammar, History, Geography, Nature Study, and Health.

Write for further information and prices

PUBLIC SCHOOL PUBLISHING COMPANY
BLOOMINGTON, ILLINOIS

Music for Christmas and New Year

Free Playlets for elementary grades

Myers. Unexpected Guests
McElroy. On Christmas Eve
Moffat. The Nativity Story
McElroy. The Little New Year
Marshall. A Toy Pageant.

Song Books

Dann. Christmas Carols and Hymns (7-8-9 Grades)	$0.80
Bryant. Christmas Carols (For Primary & Grammar Grades)	.16
Bryant. Songs for Children (4-5 Grades)	.64
Dann. Junior Songs (7-8-9 Gr.)	1.00
Johnson. Songs Everyone Should Know (7-8-9 Grades)	.72

Foresman. Books of Songs:

A Child's Book (2)	$0.52
First Book (3)	.60
Second Book (4)	.60
Third Book (5)	.64
Fourth Book (6)	.64
Fifth Book (7-8 Grades)	.88
Higher Book (H. S.)	1.52

For School Orchestras

Dann. Orchestra Parts (15 separate parts) for "Junior Songs" each (7-8-9 Grades)	.60
Dann. Conductor's Score	5.00

Write for the Free Playlets and for information concerning the textbooks

AMERICAN BOOK COMPANY

330 East 22d Street — CHICAGO, ILLINOIS

New York — Cincinnati — Chicago — Boston — Atlanta

Thousands of Successful Teachers Use "The Supreme Authority"

To get accurate, encyclopedic, up-to-date information of all kinds that is of vital use and interest in the schoolroom.

A wise school superintendent has said: *"I have never yet seen a person, whether pupil or teacher, who was accustomed to the frequent use of the dictionary who was not at the same time a good or superior all-round scholar."* A better test than this of the value of dictionary work could not be found.

The New International is constantly revised and improved to keep abreast of modern needs and information. 452,000 Entries, including thousands of New Words. 6000 Illustrations. 2,700 Pages.

Write for Helps in Teaching the Dictionary, FREE.

G. & C. MERRIAM CO.

SPRINGFIELD, MASSACHUSETTS